

From: Mertz, Prema
Sent: Friday, November 08, 2002 4:39 PM
To: STIC-Biotech/ChemLib
Subject: 09/851,595

Please search SEQ ID NO: 5 and 8 with protein databases.

Thanks

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

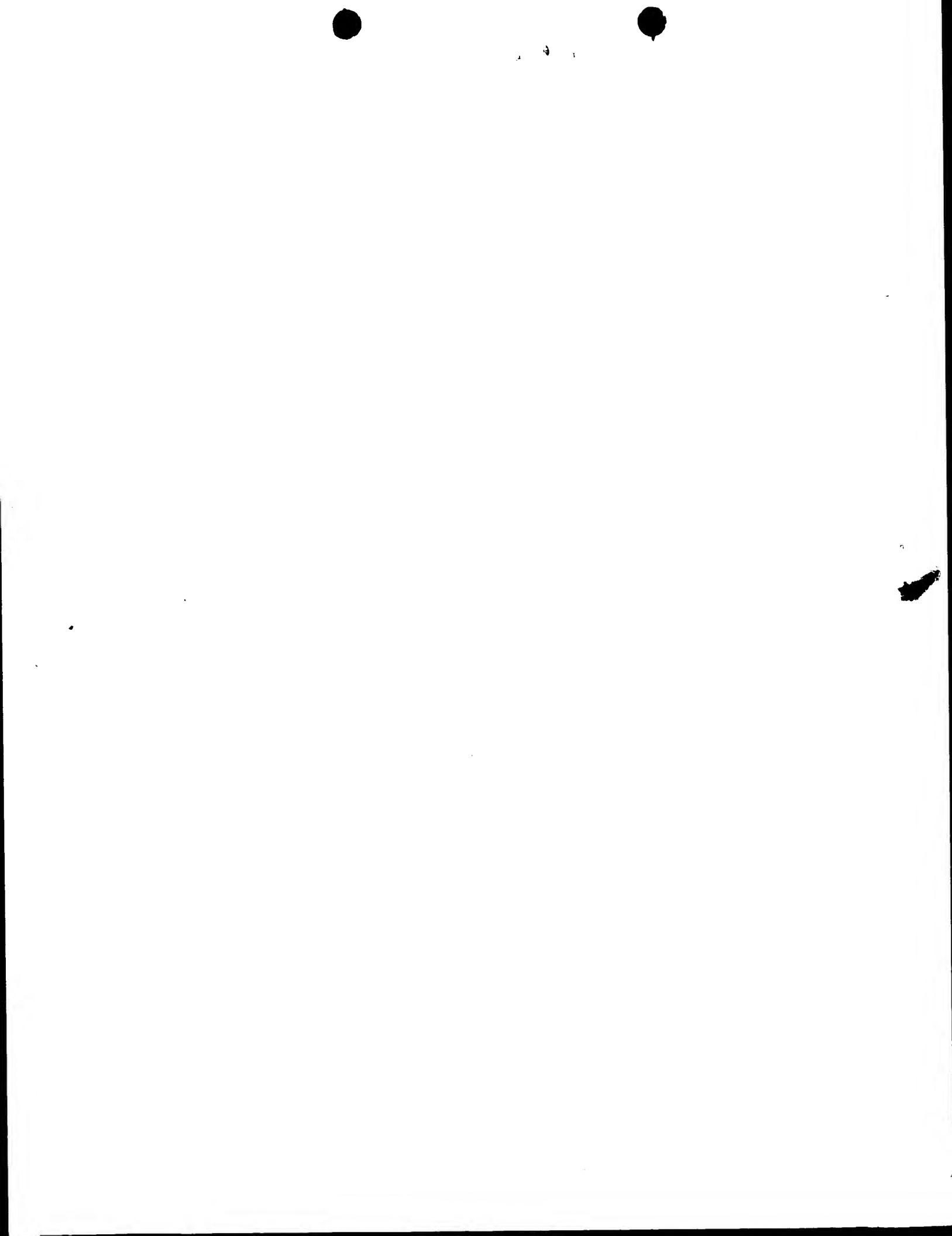
RECEIVED
NOV - 8 2002
STIC-Biotech/ChemLib
(STIC)

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 11/10/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:50 ; Search time 31.4419 Seconds
(without alignments)
4148.213 Million cell updates/sec

Title: US-09-851-595-5
Perfect score: 3307
Sequence: 1 NTHHYRESWYACRYRSGIPG.....GGLSGGGRQPSGLAFASHW 633

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rvirus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3015	91.2	928	4	Q9BYD7
2	2087	63.1	459	11	Q8R301
3	623.5	18.9	1050	5	Q9BN18
4	623.5	18.9	1360	5	Q9ND11
5	608.5	18.4	1300	5	Q9NKD6
6	549	16.6	1012	5	Q9SY16
7	549	16.6	1280	5	Q9SY17
8	459	13.9	701	13	Q9DGC6
9	439	13.3	724	13	Q9VP00
10	432	13.1	779	13	Q918M7
11	428.5	13.0	778	13	Q98TF4
12	418	12.6	658	13	Q9PVN9
13	418	12.6	763	6	Q9BGN4
14	414.5	12.5	693	13	Q9DCC5
15	414	12.5	688	11	Q64183
16	413.5	12.5	831	5	Q94979

17	412	12.5	601	13	Q42500
18	410	12.4	829	5	Q9VE64
19	408	12.3	701	6	Q8SP58
20	408	12.3	764	11	Q9D697
21	407.5	12.3	831	5	Q8SX01
22	397.5	12.0	793	13	Q91948
23	397	12.0	693	11	Q8R428
24	396	12.0	814	13	Q91949
25	392.5	11.9	662	13	Q9PW16
26	389.5	11.8	739	6	Q9B655
27	389.5	11.8	764	6	Q9B656
28	389.5	11.8	764	6	Q8SP99
29	382.5	11.6	696	13	Q98T85
30	382	11.6	533	4	Q9P1V4
31	381	11.5	699	4	Q15956
32	381	11.5	701	4	Q14751
33	380.5	11.5	696	13	Q9DGF5
34	376	11.4	662	13	Q98T84
35	371.5	11.2	470	4	Q43200
36	366.5	11.1	470	4	Q16225
37	364.5	11.0	929	5	Q18759
38	311	9.4	293	6	Q95MF7
39	291.5	8.8	673	13	Q90WP8
40	284	8.6	335	6	Q46387
41	274	8.3	590	6	Q27986
42	270.5	8.2	307	6	Q95L59
43	259.5	7.8	228	11	Q9ES08
44	255.5	7.7	307	13	Q90XC7
45	181.5	5.5	334	5	Q9VYGO

ALIGNMENTS

RESULT 1	
Q9BYD7	PRELIMINARY: PRT: 928 AA.
AC Q9BYD7	01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE VTS20631 (Fragment).	
GN VTS20631.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Euthera; Primates; Carnivora; Hominoidea; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RA Okazaki H., Hayashi A., Kozuma S., Saito T.;	
RT "A member of g-protein coupled receptor family."	
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AB049405; BAB39854.1; -	
DR HSSP; P23945; 1XUN.	
DR InterPro: IPR000276; GPCR_Rhodopsn.	
DR InterPro: IPR001611; LRR.	
DR InterPro: IPR003582; LRR-out.	
DR InterPro: IPR003591; LRR-tyr.	
DR Pfam; PF00560; LRR_13.	
DR PRINTS; PR00237; GPCR_Rhodopsn.	
DR PRINTS; PR00019; LEURICHRPT.	
DR SMART; SM00370; LRR_9.	
DR SMART; SM00369; LRR_TYR_14.	
FT NON_TER	
SQ	SEQUENCE 928 AA: 100487 MW; 4C3364ADEA89C463 CRC64;
Query Match	91.2%; Score 3015; DB 4; Length 928;
Best local Similarity	99.3%; Pred. No. 1.8e-222;
Matches 576; Conservative 2; Mismatches 2; Indels 0; Gaps 0;	
OY 54 KLEKDFPSQLSLQALDLSNAIRSIHPEAFSTLHSLVKLDTLDNQLTTLPLAGLGIMH 113	
DB 349 EIGADTFPSQLSLQALDLSNAIRSIHPEAFSTLHSLVKLDTLDNQLTTLPLAGLGIMH 408	

ID	Q9BN18	PRELIMINARY	PRT	1050 AA.
Db	121	FAEFGAAWESGIQGGQATGLAVLIGSEASVLLTLTAAVQCSTSVYCVARYKAPSPGSVRA	180	
Qy	355	GVLGCLALAGLAAALPLASVGEIGASPLCLPYAPBEGOPALGFTVALVMNSFCFLVVA	414	
Db	181	GALGCLALAGLAAALPLASVGEIGASPLCLPYAPBEGOPALGFTVALVMNSFCFLVVA	240	
Qy	415	GAVYIKLYCDLPDRGDFEAWMPCAMVRHVAWLLFADGLLYCVAFLSFSMGLPVPVEAV	474	
Db	241	GAVYIKLYCDLPDRGDFEAWMPCAMVRHVAWLLFADGLLYCVAFLSFSMGLPVPVEAV	300	
Qy	475	KSVLLVYVLPPLACLNPLLYLLENDFHFRDILRLRPAGDSGPLAYAAAGLEKSSCDSTQ	534	
Db	301	KSVLLVYVLPPLACLNPLLYLLENDFHFRDILRLRPAGDSGPLAYAAAGLEKSSCDSTQ	360	
Qy	535	ALVAFSDVDLITLASEAGRPGLLETYGEFSPVTLISQCPAPRLSESHCEPENGHGNP	594	
Db	361	ALVAFSDVDLITLASEAGRPGLLETYGEFSPVTLISQCPAPRLSESHCEPENGHGNP	420	
Qy	595	QPSMDGELLRLAEGSTPPAGGSLSGGGGFPQSLAFASHV	633	
Db	421	QPSMDGELLRLAEGATLACGGSSMGALMPSSGLFASHL	459	
RESULT 3				
ID	Q9BN18	PRELIMINARY	PRT	1050 AA.
AC	Q9BN18			
DT	01-JUN-2001	(TREMBLrel, 17, Created)		
DT	01-JUN-2001	(TREMBLrel, 17, Last sequence update)		
DT	01-MAR-2002	(TREMBLrel, 20, Last annotation update)		
DE	Leucine-rich repeat-containing G protein-coupled receptor 2.			
GN	RK OR CG8930.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID-7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Nishi S., Hsu S.Y., Zell K., Hsueh A.J.;			
RT	"Characterization of two fly LGR (leucine-rich repeat-containing G			
RT	protein-coupled receptor) proteins homologous to vertebrate			
RT	glycoprotein hormone receptors: constitutively activation of wild type			
RT	fly LGR1 but not LGR2 in transfected mammalian cells ";			
RL	Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF274591; AAK00808.1; --			
DR	Flybase: FBgn0003255; TK.			
DR	InterPro: IPR000276; GPCR_Rhodpsn.			
DR	InterPro: IPR001611; LRR.			
DR	InterPro: IPR003592; LRR_out.			
DR	InterPro: IPR003591; LRR_typ.			
DR	Pfam: PF000001; Tm_1; 1.			
DR	Pfam: PF00560; LRR_10.			
DR	PRINTS: PR00237; GPCRHHODPSN.			
DR	SMART: SM00370; LRR_8.			
DR	SMART: SM00369; LRR_typ; 10.			
DR	PROSITE: PS50262; G_PROTEIN_RECIP_F1_2; 1.			
DR	RECEPT.			

Query	March	Similarity	18.9%	Score	623.5	DB	5	Length	1050
Best	Local	Similarity	29.1%	Fred.	No.	3.8e-39			
Matches	174	Conservative	89	Mismatches	234	Indels	101	Gaps	16

QY	55	LEXPFSQSLSSLOADADLDSMNAIRSHPEAFSLHSLVKLDLIDNQLTLLPLAGLGGLMHL	114
DB	310	LPOAFQGITPKLIQLDLIDLGNEISYIHKAESGFTALIEDLNGNNLFPPELPESGLRALDL	369
QY	115	KLKGNLALSQAFKSDSEFKRLILEYPAVYOC---PYGMCASFEFAS-----	158
DB	370	KTFNNPKRIERPPDPYTFPRIGTLILSYAHCCALFLPYAMSSQKKTISOYQZEVLPFSPAE	429
QY	159	-----GOMEADLLDLDESSKRPGLGLIARQENHY	189


```

Db 430 EDTTLMNNSMMNIWPMQHNLSKOLGASHMDPWETA- INFNEQLOTOTGGQATATSYMEY 488
QY 190 DODLDE-----LQEMEDSKPHPSVOCSPTPGPKPCETLFEESGIRLAVW 235
Db 489 FEEDHVSGRATGYGFGTGLFSGMSTEDFQ- GSVQCLPMGPFPLPCADLEFDMWTLKCGW 547
QY 236 AIVLSVLCNGVLLVVFAGGAPRLPPVKEVGAIGANTLTGISCGLASVDALTFQGF 295
Db 548 VFLLSLGNGVVFVLLC-SRSKMDVPRFLVGNLAADPFMGITGLAIYDAATLGEF 606
QY 296 SEYGARETGLGRATGFLAVLGSSEASVLLTLTAOVCSVSCVARYGKSPSLGVRAG 355
Db 607 RMFALIPWQMSVLCOLSGFLAVLSSELSV--YTLAVITLERNVAITHAHLNKRSLKQAG 664
QY 356 VLGL--ALAGLAALPLASVGEYASPLCPYAPREGOPALGFVALVMNSFCFLVY 413
Db 665 YIMSVGWVFALIMLMPVGVSDYRKFAVCLPEFTTG-PASLTYISLMFINGCAFLTL 723
QY 414 AGAYIKLYCDLPRGDFEAVW---DCAMVRHVAWLIIFADGLLYCVAFLSFASMLGFPVT 470
Db 724 MGCYLMKWAI-RG--SOAMNTNDSRIAKRMALLVFTDFLCWSPIDAFESITAIIFGLQLIS 780
QY 471 PEAVKSVLLVPLPACINPLLYLLENFPERDRLRLPR-----AGDSGPLAYAAAGE 524
Db 781 LEQAKITFVFLPLNMSCNPFPLVIMTKQFKDCVTLCKHFEESRVYGGGPGRGAVAR 840
QY 525 LEKSCDSTQALVAFSDVLDL---ILEASEAGRPGLTYGPPSVTLISCOQPGAPRLE 579
Db 841 TKRG-----DLPPPLPAAVAHPPGCR-----CLRMLPSEMPNMHKKME 879

RESULT 4
Q9NDI1 PRELIMINARY: PRT: 1360 AA.
AC 09NDI1:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Glycoprotein hormone receptor II.
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CANTON S; TISSUE=WHOLE ANIMAL;
RX MEDLINE=20359836; PubMed=10899142;
RA Erikson K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA Grimmelshuisen C.J.P.;
RA "Molecular Cloning", Genomic Organization, Developmental Regulation,
RT and a Knock-Out Mutant of a Novel Leu-Rich Repeats-containing G
RT Protein-Coupled Receptor (Dlcr-2) from Drosophila melanogaster.";
RL Genome Res. 10:924-938(2000).
DR EMBL: AF142343; AAF6608.1;
DR HSP: O57815; ID3Y.
DR FLYBase: FBgn0003255; RK.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_tyr.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_14.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR SMART: SM00370; LRR_2.
DR SMART: SM00369; LRR_TYR_5.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
RN SEQUENCE 1360 AA; 150731 MW; 7043515B4F6612 CRC64;
SQ
Query Match 18.9%; Score 623.5; DB 5; Length 1360;

```

```

Best Local Similarity 29.1%; Pred. No. 5.2e-39;
Matches 174; Conservative 89; Mismatches 224; Indels 101; Gaps 16;

QY 55 LEQDFQSSLSQALDLSMNAIRSHPEASTLSHLYLDTLDNQLTTLPLAGIGMLH 114
Db 524 LPQDAFQGPQLQDLDEGNETSYTHKRAFSTALDELNNGNIFPLPSGIRALLHL 583
QY 115 KIKGNLALSQAESKDSFPKRLILEVPVAVQCC---PYGMCSFRRAS----- 158
Db 584 KTFNNPKLKEFPPTFPRIQTLISYAHCCAFPLVAMSSQKTSQVQEARVLPSPAE 643
QY 159 -----GWEADHLHDESSKRPLGLARQENHY 189
Db 644 EDTTLMNNSMMNIWPMQHNLSKOLGASHMDPWETA- INFNEQLOTOTGGQATATSYMEY 702
QY 190 DODLDE-----LQEMEDSKPHPSVOCSPTPGPKPCETLFEESGIRLAVW 235
Db 703 FEEDHVSGRATGYGFGTGLFSGMSTEDFQ- GSVQCLPMGPFPLPCADLEFDMWTLKCGW 761
QY 236 AIVLSVLCNGVLLVVFAGGAPRLPPVKEVGAIGANTLTGISCGLASVDALTFQGF 295
Db 762 VFLLSLGNGVVFVLLC-SRSKMDVPRFLVGNLAADPFMGITGLAIYDAATLGEF 820
QY 296 SEYGARETGLGRATGFLAVLGSSEASVLLTLTAOVCSVSCVARYGKSPSLGVRAG 355
Db 821 RMFALIPWQMSVLCOLSGFLAVLSSELSV--YTLAVITLERNVAITHAHLNKRSLKQAG 878
QY 356 VLGL--ALAGLAALPLASVGEYASPLCPYAPREGOPALGFVALVMNSFCFLVY 413
Db 879 YIMSVGWVFALIMLMPVGVSDYRKFAVCLPEFTTG-PASLTYISLMFINGCAFLTL 937
QY 414 AGAYIKLYCDLPRGDFEAVW---DCAMVRHVAWLIIFADGLLYCVAFLSFASMLGFPVT 470
Db 938 MGCYLMKWAI-RG--SOAMNTNDSRIAKRMALLVFTDFLCWSPIDAFESITAIIFGLQLIS 994
QY 471 PEAVKSVLLVPLPACINPLLYLLENFPERDRLRLPR-----AGDSGPLAYAAAGE 524
Db 995 LEQAKITFVFLPLNMSCNPFPLVIMTKQFKDCVTLCKHFEESRVYGGGPGRGAVAR 1054
QY 525 LEKSCDSTQALVAFSDVLDL---ILEASEAGRPGLTYGPPSVTLISCOQPGAPRLE 579
Db 1055 TKRG-----DLPPPLPAAVAHPPGCR-----CLRMLPSEMPNMHKKME 1093

RESULT 5
Q9NKD6 PRELIMINARY: PRT: 1300 AA.
AC 09NKD6; Q9YJ03;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 144.0 kDa protein (Rk gene product).
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RX MEDLINE=9403001; PubMed=10471707;
RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazek R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Gelinkner S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
SQ

```

RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
 RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummel S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomoth M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacle J.M., Park S., Pfeiffer B., Poon L., Sequiera A.,
 RA Seeth H., Smit E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieren L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
 RP STRAIN-BERKLEY.

RA MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Wan K.H., Doyle C., Baxter E.G., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Basu A., Bendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Burtis K.C., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertile W.S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hattis N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwu C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mount S.M., Milshina N.V., Mobarry C., Morris D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Port J., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03408; AAF44846.1; -;
 DR EMBL: AEO03642; AAF53367.2; -;
 DR FLYBase: FBgn0003255; rk-;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_Typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 10.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR SMART: SM00370; LRR; 1.
 DR SMART: SM00369; LRR_Typ; 4.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR KW Hypothetical protein.
 SO SEQUENCE 1300 AA; 144031 MW; B4B9E39F942FA0B3 CRC64;

Query Match 18.4%; Score 608.5; DB 5; Length 1300;
 Best Local Similarity 28.6%; Pred. No. 6.9e-38;

Matches 174; Conservative 89; Mismatches 234; Indels 111; Gaps 17;

OY 55 LEXDFFSLSLQAL-----DLSMNAIRSIHPEAFSLHSLVRLDLDNDQLTTP 104

Db 454 LPDAFQGIPIKQLQLVYSITRIHSLEGNISYTHKEAFSGTALDELNLGNIFPELP 513
 OY 105 LAGGIMHLMLKGNLALSOAFSKDSFPRKRLIEVPVYVOC---PYMCASFRRAS--- 158
 Db 514 ESGRALLHLTKFNPNKLRPEPPDPFFRIQTLISYVHCCAFILPVAMSSOKRTSOVQ 573
 OY 159 -----GQMEADHLIDDESSKRPGL 179
 Db 574 EAVLFPDAEDPDKLNNMNMNIMPOMHNLKQLGASMHDPWETA-INFNEDQIQDTGG 632
 OY 180 LLARQENHYDDLDDE-----LQLEMSKRPSPVOCSPTRPEPCETLF 225
 Db 633 QIATSYMEEFEEHEDVGPAATGFGGLSGMSTEDPQ-GSVQCPLPPEPPLCADLF 691
 OY 226 ESMGIRLAVNAIVLISLVCGLVLTIFAGGPADLPVKKFVGAIGANTLTGISCGLA 285
 Db 692 DWMTLRCGVVWVFLSLTGLGTVFVLLC-SRSKMDPRFLVNCMAADPFMGIYGLA 750
 OY 286 SVDAITFGQFSEYGARWETLGRATGFLAVLSEASVLLTLTAAGCSVSCVARYGK 345
 Db 751 IYDAATLGEFRMFALPQMSVLQGLSFLAVLSLSV-YTLAVITLERVYATTHAHL 808
 OY 346 SPGLSVRAGVGLCL--ALAGLAALPLASVGEYASPLCLPYAPPEGCPALGTTVALY 403
 Db 809 NKRLSLKQGYIWSVGFALMALMPVGVSDYRRFAVCLPETTTG-PASLTGYISLM 867
 OY 404 MNNSFCELVAGAVIYLCYDLPGRDFEAVV--DCAMVHVAVMLIPADGLICPAFLSE 460
 Db 868 FINGCAFLLMGCYLKMVAI-RG--SQAMNNDSTAKRMALVYTDLCSPIAFFSI 924
 OY 461 ASMLGFPPTPEAVSVLVLPPLACNPLLYLFPNFRDLRLRPR-----AGDS 514
 Db 925 TAIFFQLTSLDQATFTVFLPLNSCCNPFLYATMTQFKKDCVYLCKHPESSRYVGG 984
 OY 515 GPLAAMAAGELKSSCDSTQALVAFSDVL--ILEASAGRPPLGTYTFPSVLLISQ 571
 Db 985 GGGCGAVARTRG-----DLPPPLPAVAHPPCR-----CLMRPSE 1025
 OY 572 GPGARPLE 579
 Db 1026 MPMNHKME 1033

RESULT 6
 ID 095Y16 PRELIMINARY; PRT; 1012 AA.
 AC 095Y16
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, last annotation update)
 DE Glycoprotein hormone receptor.
 GN ACPHPR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Asterozoa; Valvatacea; Valvatida; Asterinidae; Asterina.
 OX NCBI_TaxId=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oda Y., Yoshikuni M., Nagahama Y.;
 RT "CDNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish *Asterina
 pectinifera*.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB061862; BAB68209.1; -;
 DR InterPro: IPR002106; ALCRNA_LigaseII.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR
 DR InterPro: IPR00372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 14.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PRINTS; PR00237; GPCR RHODOPSIN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; UNKNOWN_1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE 701 AA; 78453 MW; B3D78465CA5410A CRC64;

Query Match 13.9%; Score 459; DB 13; Length 701;
 Best Local Similarity 25.6%; Pred. No. 9.2e-27;
 Matches 149; Conservative 85; Mismatches 241; Indels 106; Gaps 15;

```

QY 16 SGIGSTHVASERSOGSLPAH-----ASLALASANT-----TASGKLE 56
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 131 SNDELSEIITTKSHLSF-IHPDAFRMARLREPLTISNTGLRIFFPDRSKTHSTACFLD 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 XDFESQSLSDA-----LDLSMAIRS-----I 79
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 LQDSHTRKVPANAFRLCTOTFAEIRLTRNGIKVEASDAFNGTKMHLFEGCNBOLTHI 249
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 80 HPEAFSTLHSLVKLDLTDNOLTTPLAGLGIMHLKLGKLNALSQAESKDSFPKRLILEV 139
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 250 SPNAFVGSSELYLVLDVSEETALTSIPDSILDGLKRLIAESAFNKEPPIQLFTLHQAKL 309
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 140 PYAVOCPPYGCASFEEKASQWEADLHLDDESSKRPGLLAQAEHNTDQDLEOLE 199
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 TYSHCCAF---LNMHRNSRMS---LCDNPEAKNNLHFREYCSN----- 350
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 MEDSKPHSVQCSPTPGPFKPCETLPESWGIRLAWAIVLVLSCNGLVLTVPAGGPAP 259
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 351 -----STNTCSBPAPDFNCECDIMSATPLRLIWIISVALLGNVLL-VLLGSRYK 403
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 LPPVKFVVGAIAGANTLTGISGLASVDALTFQSESEYCARWETGLGCRATGLAVLGS 319
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 404 LTPVRLMCLAPADLCMGYLYLVATVDMLTRGRYNYAIDQMOMGICNAAGFEYFAS 463
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 320 EASVLLLTAAVOCVSVSCVRAVYKSPSLGSVAGVLCAL-----AGLAAALPLAS 373
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 464 ELISVFLTATVBRMHTITHALDLRLRHR-----ACIIMTIGISFLAALLFTVG 517
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 374 VGEYGSPLCLPAPPEGOPALGFTVALVMNMFCEFLVAVAGAYIKLYCDLPRGDEFAVW 433
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 518 ISSYGVISICLPW-DVESLVSQFYVYVCLLNLILAFVCCGYLSYLFERRPSSAAH 575
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 434 -CCAMVRHVAWLFADGLICPAVFLSPASMLGFPVTPRAVSVLLVPLACLPIL 492
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 576 AOTRVAQRAVLIPTFCIMAPISFAISPAALKPLITVSDSKLLLVLEVPINSCSNPL 635
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 493 YLLENPHERDILRLRPRAGDSGPLAYAAAGEL---EKSSC 530
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 636 YAFETNFRERDFELLAARF-----LFKTRAQIYRTRESSC 671
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 9

Q9PVP0 PRELIMINARY; PRT; 724 AA.
 AC Q9PVP0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Gonadotropin receptor II.
 GN SGRH-RII.
 OS Oncorhynchus rhodurus (amago).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=41164;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94423507; PubMed=10491336;
 RA Oba Y., Hirai T., Yoshura Y., Yoshikuni M., Kawachi H., Nagahama Y.,
 RT "Cloning, functional characterization, and expression of a
 RT gonadotropin receptor cDNA in the ovary and testis of amago salmon
 RT (Oncorhynchus rhodurus)."
 RL Blochem. Biophys. Res. Commun. 263:584-590(1999).

DR EMBL; AB030005; BAA84638.1; -
 DR HSSD; P22888; ILOT.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR000372; LRR_Nterm.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR RHODOPSIN.
 DR SMART; SM00013; LRRT; 1.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; UNKNOWN_1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE 724 AA; 80494 MW; E999652740332B2E CRC64;

Query Match 13.3%; Score 439; DB 13; Length 724;
 Best Local Similarity 26.7%; Pred. No. 3.3e-25;
 Matches 125; Conservative 84; Mismatches 210; Indels 50; Gaps 9;

```

QY 76 IRSIHPEAFSTLHSLVKLDLTDNOLTTPLAGLGIMHLKLGKLNALSQAESKDSFPKRL 135
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 218 LRVTHREAFKAVGRPRLIDVSTAIETPLSHGLSVVELVARTAYGLKRLPPFRDLGNLQ 277
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 136 ILEVPIAYQCCPYGCASFEEKASQWE-----AEDLHLDDESSKRPGLIL 181
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 278 KAHLTYSNHC-----ALLTWDTHRDSPINEAQHNGSRPTYCDDSPSKFPAGMV 327
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 -----ARQAEHNTDQDLEOLEMEDSKRHPVQCSPTPGPFKPCETLPESW 228
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 DSSDTSLIVEIHGTNEBSGVGVPELWL-NCQTRPTLQCPREDAFNPCDIAFGS 386
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 229 GIRLAWAIVLVLSCNGLVLTVPAGGPAPLPVKFVVGAIAGANTLTGISGLASVD 288
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 387 FLRAVWFNLTALIGN-LTVLLIFTSRCKLTPRRLMCLHAFADICIGYLLMTAIVD 445
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 289 ALTFPGQSESGARKRETGLGCRATGLAVLGSASVLLTLTAANO---CSVSVCAVAY 344
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 446 LHTRGHSEHAIDWQTAGSAGFELSGELSVYLTITLERMHTITHALEKRLG 505
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 345 KSPSLGSVRAGVLCALAGLAAALPLASVGEASPLCLPYAPPEGOPALGFTVALVM 404
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 506 LAQAGIMAGWILICM-----AMPLVGVSYSTRVSMCLPM-DVKTPLAQAFILLL 559
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 405 MNSFCPLVAVAGAYIKLYCDLPRGDFEA-VWDCAMVRHVAWLFADGLICPAVFLSPAS 463
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 560 FNVGAEFLWCVCYVLYLAVARNQFPFSRSADAKIAKMAVLIPTDLCMAPIFFPISAA 619
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 464 LGLEPVTPEAVKSVLLVPLPACLPILYLLFNPHFRDLRLRPRAG 512
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 620 FKVPLITVNSKILLVLPFINSCLAMPFLYAITKAFKRDVYLLLSNIG 668
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 10

Q9I8N7 PRELIMINARY; PRT; 779 AA.
 AC Q9I8N7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Thyrotropin receptor precursor
 OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=20457109; PubMed=11000515;
 RA Kumar R.S., Jilka S., Kight K., Swanson P., Diltman A., Alok D.,
 RA Zohar Y., Trant J.M.,
 RT "Cloning and functional expression of a thyrotropin receptor from the
 RT gonads of a vertebrate (bony fish): potential thyroid-independent role
 RT for thyrotropin in reproduction."
 RL Mol. Cell. Endocrinol. 167:1-9(2000).

[illegible][illegible]

Db 670 LVLVFLPLNSCANPFLVAFITKAFRGDIFLLSKVG 705

RESULT 12

Q9PVN9 PRELIMINARY; PRT; 658 AA.
 AC Q9PVN9;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Gonadotropin receptor I.
 GN SGNH-RI.
 OS Oncorhynchus rhodurus (amago).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=41164;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20025349; PubMed=10558873;
 RA Oba Y., Hirai T., Yoshida Y., Yoshikuni M., Kawachi H., Nagahama Y.,
 RT "The duality of fish gonadotropin receptors: cloning and functional
 RT characterization of a second gonadotropin receptor cDNA expressed in
 RL the ovary and testis of amago salmon (Oncorhynchus rhodurus).";
 RL Biochem. Biophys. Res. Commun. 285:366-371(1999).
 DR EMBL: AB030012; BAA8698.1; -.
 DR HSSP: P22888; 1LUT.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_2.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 658 AA; 72991 MW; 853A38BEA845480 CRC64;

Query Match 12.6%; Score 418; DB 13; Length 658;
 Best Local Similarity 24.7%; Pred. No. 1.2e-23;
 Matches 139; Conservative 84; Mismatches 238; Indels 102; Gaps 10;

QY 16 SGIPGTHASVERSGSLPAHPASIALAASNTTA--SGLENDTFEQLSS----- 65
 Db 98 ANPRLTEITTSKHLVITHQAFILPKLSHLTICHTGRLVLPNFSRISHAAMTFLD 157
 QY 66 -----LQALDLSMNAIRSIHEAF--STLSLVK----- 92
 Db 158 LQDNVHVIIPSNALGLTNTIDELRLTKNGISEVESHAENGTKIRKILMGNQLSHM 217
 QY 93 -----LDLTDNQLTTPPLAGLGMLKLKGNLALSQAESKDSFPKRLILEV 139
 Db 218 HNSSEKGAEGPFLDISRTALSSLPESVLGEVHLSAVSVFSLRPLPLSLFTKRLQANL 277
 QY 140 PYAVQCC-----PYGMCASFEEKASGOWEAEEDLHLDDESSKRPGLLQAEENYD 190
 Db 278 TVPSHCCAFHKHQNRNFRFTMSACFKGAQ----- 307
 QY 191 QDDELQLEMEDESKRHPVSQCCSPPPGPKCEYLPESWGIRLAWAIVLLSVCNGVLV 250
 Db 308 ---NNHAFPEDFCLMNTSVACSPAPDAFNPCEIDIMGSAPLRLIITISVLLALGNTIVLL 364
 QY 251 TVPAGGPAPLPYKFFVGAAGANTLTGISCGLASVDALTFGQFSEVGARWETGLGCR 310
 Db 365 -VLLSGRAKKTVPRLMLCHLSEFADLCMGITLVATVDVTRGLYVNIATISWOTGAGCDI 423
 QY 311 TGFVLAVGSEASVLLTLAAVQSVSVSCVAVKSPSLGSRAGVLCCLALACLAALP 370
 Db 424 AGFFTVFASBELSMFTLTATLTLEXYHTTHALRLDRKLRLRACAVMTGAFSCLAALLP 483
 QY 371 LASVGYGASPLCLPYAPPGOPALAGFTYALVWMSFCELVAAGAVIKLYCDLPGRGE- 429
 Db 484 TVGVSSYSKVSICLP-MVESLPSOV-FVMEFLLLNVAVFLCVCLISTYLVSRRNSSP 541

QY 430 EAVWDCAVNRVAVLITADGILLYCPVAFLSFASMLGLFPVTPPEAVKSVLVVLPACIN 489
 Db 542 PASAETCMQAFMALITITDFLCMAPISEFFALSAALIKLPLIVSDSKLLVLYFPINSCAN 601
 QY 490 PLVLYLFNPHRPDDLRLRLPRAG 512
 Db 602 PFLYGLCTRTFRDRDFLLAARYG 624

RESULT 13

Q9BGN4 PRELIMINARY; PRT; 763 AA.
 ID Q9BGN4;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Thyrotropin receptor.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen L.Q., Karamanoglu Arseven O., Gerber H., Stein B.S.,
 RA Jameson J.L., Kopp P.;
 RT "Cloning of the cat thyrotropin receptor and evidence against an
 RT autoimmune etiology of feline hyperthyroidism";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF218264; AAK00133.1; -.
 DR HSSP: P16473; 1XDM.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 763 AA; 86588 MW; 937F60A140487D60 CRC64;

Query Match 12.6%; Score 418; DB 6; Length 763;
 Best Local Similarity 24.4%; Pred. No. 1.4e-23;
 Matches 144; Conservative 108; Mismatches 254; Indels 84; Gaps 15;

QY 33 SLPAHPASIALAASNTTASGKLENDTFEQLSS-----SSLDLDSMNA-IRSIHEAF 84
 Db 166 SIPAN-----AFQGLCEMETLTLKLYKNGFTSIGQAHAFNGTKLDAVYLKNNKYLITADDAF 221
 QY 85 STLHS-LVKLDLTDNQLTTPPLAGLGMLKLKGNLALSQAESKDSFPKRLILEVYAY 143
 Db 222 GGVSQGPLTLDVSYTSVTLPSKGLHKLKLIRNMTWTLKPLTLISFLHUTRADLSYPS 281
 QY 144 OCCPY-----GMCASF-----KASGWEAE-----DLHLDDESS 174
 Db 282 HCCAFNRNOKKIRKILDESFCNDSSIRSLRQKSVNALNGFDEYEBEYGLDSHAGYDNS 341
 QY 175 KRPLGLARQAEHNY-----DQDL-----DELQLEMEDESKR-----HPSVOC 211
 Db 342 K-----FQDTRNSHYVVFEEDEILGFGQELKNPQETLQAFDSHYDYTVCGNENEMVC 397
 QY 212 SPTPGPKCEYLPESWGIRLAWAIVLLSVCNGVLVLTVPAGGAPLPYKFFVGAIA 271
 Db 398 TPKSDEFNPCEIDIMGYKPLVYVWFSLLALGNVFLVLLITLSHYKLPV-RELKCNLA 456
 QY 272 GANTLTGISCGLASVDALTFGQFSEVGARWETGLGCRATGFLAVGSEASVLLTLAAV 331
 Db 457 FADFCMGWTLULLIASVDLTHSEYINHAIDWOTGPCNNAAGFTTVRASLSVTLVITL 516
 QY 332 QCSVSVSCVAVKSPSLGSRAGVLCCLALAGLAALPLASVGEYGASPLCLPYAPPEG 391
 Db 517 ERWYATTFAMRLDRKMRRLRHAVAIWVGWVCCFLALLLPVGVSSYAKVSICLPM--DTE 574
 QY 392 QPALGFTYALVWMSFCELVAAGAVIKLYCDLPGRGE-AVWDCAVNRVAVLITADGL 450

```

Db 575 TPLALATIIIVLLNINIAFTIVCGCYKIITYIVANPCYNTGDKDTAKRAVALIFTEDEM 634
OY 451 LYCPVAFLEASMLGFPVTPBAVKSULLVLPACLNPLLYLLEFNHFRDDLRLRPR 510
Db 635 CMADISFALSAALMKNKPLITVTNSKILLVLEFPYNSCANPELYAIFTTFORFILLSK 694
OY 511 AGDSGSPLAYAAG-----ELEKSSCDSDGQALVAFSDVDLILBAS 549
Db 695 FGICKRQAOAVRGORVSPKSTGIQVCKVTNRNMSLPRMODDYELLNS 744

RESULT 14
ID 09DGC5 PRELIMINARY: PRT: 693 AA.
AC 09DGC5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Gonadotropin receptor II.
OS GnRH-R1.
OC Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Oba Y., Hirai T., Yoshitura Y., Yao Z., Nagahama Y.;
RT "Tilapia gonadotropin receptor II."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041763; BAB16107.1; -.
DR HSSP; P22888; ILUT.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECPEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 693 AA; 77856 MW; C3BBDPDEFCF988 CRC64;

Query Match 12.5%; Score 414.5; DB 13; Length 693;
Best Local Similarity 28.0%; Pred. No. 2.3e-23;
Matches 134; Conservative 79; Mismatches 204; Indels 61; Gaps 14;

OY 61 SOLSSIQALDLSNNAIRSHPEAFSTLHSLVKLDLTDNQLTTPPLAGLGMLHLKLGNL 120
Db 200 TKLNTLVLRD-NM-YLRNIQEDAFEGATGPTLLDVSSTALRSLP--PGLRHVKF--L 251
OY 121 ALSQASKSPFRLRL-----EVPYAYQCCPYGMCASFKASGQWEADLHIDDESS 174
Db 252 KASHAALASLPLESLAELEELTYPSHC-----AFHWRRK-----QRESA 296
OY 175 KRPLGLLAROENHYDODDELQI-----EME-DSKPHPSVOCSPRPGRPKPEYL 224
Db 297 LKNLTFKCDLMTNEIPTADDTSLINDINFOYDLEFDCSNPFVKSPKPDAFNPCEDL 356
OY 225 FESWGIRLAWMAVLVLSVLCNGVLVLFVAGCAPLPVFEVGAIGANGMLNGISGLL 284
Db 357 LGFSFRLCLTWITMVAAGN-LAVLVILLGHKRLTVSRFLKCNLAFLADLCMLYLILI 415
OY 285 ASDVALTFQFSESGARWETGLCRATGFLAVLGSSEASVLLTLAAVQCSVSVCAVAG 344
Db 416 AFDYHSHHHYNNHATDWMGPGCGAGFLVPSSELSVTLTVTSLEPMHTTNAMHVN 475
OY 345 KSPSLGSVRAVLGIALAALPLASVGEYASPLCLPYAPBPQPAALG---FTVA 401
Db 476 KRLRMHHVTAMVGMGAFSLVALLPLVGVSVYSKVSICLPM-----DIDTLGAQVYVA 530
OY 402 LVMAANSICFLVAVAGYIKIKYCDL-----PRODFEAVMPCAMVRHYAMLFADGGLYCP 454

```

```

Db 531 VLLNVAFLVVCYCYIGIYISVHNPEHSTRGRDK-----IAKRMVLIPTDPLCAP 584
OY 455 VAFLEASMLGFPVTPBAVKSULLVLPACLNPLLYLLEFNHFRDDLRLRPRAG 512
Db 585 ISFEAISALRNPILTVSHSKILLIFPINSCLNPFLYITTFRAFRKDYCLLSRCG 642

RESULT 15
ID 064183 PRELIMINARY: PRT: 688 AA.
AC 064183;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Follicle-stimulating hormone receptor.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92149579; PubMed=1738373;
RA Heckert L.V., Daley I.J., Griswold M.D.;
RT "Structural organization of the follicle-stimulating hormone receptor
RT gene."
RL Mol. Endocrinol. 6:70-80(1992).
DR EMBL; S81198; AAB21415.2; -.
DR EMBL; S81177; AAB21415.2; JOINED.
DR EMBL; S81179; AAB21415.2; JOINED.
DR EMBL; S81171; AAB21415.2; JOINED.
DR EMBL; S81174; AAB21415.2; JOINED.
DR EMBL; S81178; AAB21415.2; JOINED.
DR EMBL; S81183; AAB21415.2; JOINED.
DR EMBL; S81185; AAB21415.2; JOINED.
DR EMBL; S81194; AAB21415.2; JOINED.
DR HSSP; P23945; LXUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECPEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 688 AA; 77341 MW; 441F0D9E7D01DF18 CRC64;

Query Match 12.5%; Score 414; DB 11; Length 688;
Best Local Similarity 24.8%; Pred. No. 2.5e-23;
Matches 165; Conservative 86; Mismatches 253; Indels 160; Gaps 24;

OY 16 SGIRGSTHNASVERSGQL-----SLPAHPASTIAIAASNTTASGLKEXDFQSLSL 66
Db 92 SNLPKLEIRKRNKNNLLYNPEAFQNLK-----SLRVLISNT-----GIKHLPAVHKIOSL 144
OY 67 Q-----ALDSM-----NAIRSHPEAF-----STLHS 89
Db 145 QKVLDDIQDNINIHIVARNSEFMGLSFEMLSKNGIEEIHNCAFNGTQDELNLSNNNIEE 204
OY 90 L-----VKLDLTDNQLTTPPLAGLGMLHLKLGMLALSQAQSKSPFRLRLTE 138
Db 205 LPNDVFGASPVLLDLSKRYVHSLPNHGLENLKRLARSYRMRKLPNDLKFTYLMRAS 264
OY 139 VPYAYQCCPYGMCASFKASGQWEADLH-----LDDDES 173
Db 265 LTPSHCCAF-----ANLKRQISELHPICNKSILRQDIDMTQIGDQVSLIDEPSS 316
OY 174 SKRPLGLLAROENHYDODDELQIEMDSKPHPSVOCSPRPGRPKPEYLFEESWGIRLA 233
Db 317 ---YKGSDDMMYNEFDYLCN---EYVD-----VTCSPKPDANPCEDIMGINILRYL 363

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:31:55 ; Search time 9.24763 Seconds
(without alignments)
2839.055 Million cell updates/sec

Title: US-09-851-595-5

Perfect score: 3307
Sequence: 1 NPTTHRESWYACRYSRGPIC.....GGLSGGGGRQPSGLAFASHV 633

Scoring table: BLASTSUM62
Gapop 10.0 , Capext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3011	91.0	828	1	LGR6_HUMAN
2	1275	38.6	907	1	LGR5_HUMAN
3	1226	37.1	907	1	LGR5_MOUSE
4	967	29.2	951	1	LGR4_RAT
5	957	28.9	951	1	LGR4_HUMAN
6	440	13.3	692	1	FSHR_HUMAN
7	428.5	13.0	694	1	FSHR_HORSE
8	424.5	12.8	695	1	FSHR_BOVIN
9	422	12.8	695	1	FSHR_PIG
10	419	12.7	696	1	FSHR_SHEEP
11	418.5	12.7	695	1	FSHR_RAT
12	417	12.6	692	1	FSHR_HUMAN
13	414.5	12.5	695	1	FSHR_SHEEP
14	414.5	12.5	695	1	FSHR_MACACA
15	414.5	12.5	764	1	FSHR_CANINA
16	414.5	12.5	764	1	FSHR_SHEEP
17	412.5	12.5	700	1	FSHR_MOUSE
18	409	12.4	763	1	FSHR_BOVIN
19	408	12.3	764	1	FSHR_MOUSE
20	407.5	12.3	700	1	FSHR_RAT
21	404	12.2	701	1	FSHR_BOVIN
22	402	12.2	687	1	FSHR_EQUUS
23	398.5	12.1	764	1	FSHR_HUMAN
24	381.5	11.5	764	1	FSHR_RAT
25	381	11.4	699	1	FSHR_HUMAN
26	377.5	11.4	693	1	FSHR_CHICK
27	377	11.4	676	1	FSHR_CALA
28	363	11.0	538	1	FSHR_SHEEP
29	357	10.8	366	1	FSHR_CHICK
30	309	9.3	754	1	LGR8_HUMAN
31	308.5	9.3	925	1	GLHR_ANCEL
32	306.5	9.3	737	1	LGR8_MOUSE
33	299	9.0	1115	1	GPCR_LYMST

34	269.5	8.1	757	1	LGR7_HUMAN	O9hbx9 homo sapien
35	149.5	4.5	536	1	CBP8_HUMAN	E22792 homo sapien
36	138	4.2	646	1	FLR1_HUMAN	Q9uzul homo sapien
37	137.5	4.2	355	1	OPSB_GECGE	P35357 gecko gecko
38	135.5	4.1	370	1	GALT_RAT	O88626 rattus norv
39	134.5	4.1	372	1	GALS_RAT	O08726 rattus norv
40	134	4.1	354	1	OPSD_GALML	O93441 galeus mela
41	134	4.1	371	1	GALS_MOUSE	O88854 mus musculu
42	133.5	4.0	370	1	GALT_MOUSE	O88853 mus musculu
43	132	4.0	377	1	APJ_MOUSE	Q9wv08 mus musculu
44	132	4.0	377	1	APJ_RAT	Q9jhg3 rattus norv
45	131	4.0	350	1	MLIA_HUMAN	P48039 homo sapien

ALIGNMENTS

RESULT 1
LGR6_HUMAN STANDARD: PRT; 828 AA.

AC Q9HbX8; Q96K69; 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

SEQUENCE FROM N.A.

RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,

van der Spek P.J., van Duin M., Hsueh A.J.W.;

"The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signaling mechanism for LGR7."

Mol. Endocrinol. 14:1257-1271(2000).

SEQUENCE OF 406-828 FROM N.A.

Takeda S., Kadowaki S., Haga T., Takeasu H., Mitaku S.;

"Identification of G protein-coupled receptor genes from the human genome sequence."

Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

SEQUENCE OF 431-828 FROM N.A.

Tissue-Mammary gland;

Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,

Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,

Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,

Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;

"NEO human cDNA sequencing project."

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

FUNCTION: Orphan receptor.

SUBCELLULAR LOCATION: Integral membrane protein.

SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

SIMILARITY: CONTRAINS 11 LEUCINE-RICH REPEATS (LRR).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

EMBL: AB083616; BAB89329.1; -

EMBL: AK027377; BAB55071.1; ALR_INIT.

EMBL: AF190501; AAC17168.1; -

```

DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_tyr.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 8.
DR PRINTS: PRO1361; EDG8RECEPTOR.
DR PRINTS: PRO0237; GPCRHOPOSPN.
DR PRINTS: PRO0019; LEURICHRPT.
DR SMART: SM00370; LRR; 5.
DR SMART: SM00369; LRR; 10.
DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; FALSE, NEG.
DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; FALSE, NEG.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
Leucine-rich repeat.
FT DOMAIN 1 428 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 429 449 1 (POTENTIAL).
FT DOMAIN 450 459 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 460 480 2 (POTENTIAL).
FT DOMAIN 481 505 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 506 526 3 (POTENTIAL).
FT DOMAIN 527 548 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 549 569 4 (POTENTIAL).
FT DOMAIN 570 588 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 589 609 5 (POTENTIAL).
FT DOMAIN 610 635 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 636 656 6 (POTENTIAL).
FT DOMAIN 657 670 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 671 691 7 (POTENTIAL).
FT DOMAIN 692 804 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 813 816 POLY-GLY.
FT DOMAIN 822 45 LRR 1.
FT REPEAT 46 69 LRR 2.
FT REPEAT 71 93 LRR 3.
FT REPEAT 94 117 LRR 4.
FT REPEAT 118 140 LRR 5.
FT REPEAT 142 164 LRR 6.
FT REPEAT 168 211 LRR 7.
FT REPEAT 212 236 LRR 8.
FT REPEAT 238 257 LRR 9.
FT REPEAT 258 281 LRR 10.
FT REPEAT 283 305 LRR 11.
FT DISULFID 503 578 BY SIMILARITY.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 406 410 CSPTP -> MISPT (IN REF. 2).
FT CONFLICT 628 628 W -> R (IN REF. 3).
FT CONFLICT 824 824 FASHV -> LHHTY (IN REF. 1).
SQ SEQUENCE 828 AA; 89301 MW; 1B5971445AAD8B4 CRC64;
Query Match 91.0%; Score 3011; DB 1; Length 828;
Best Local Similarity 99.1%; Pred. No. 2,5e-214;
Matches 575; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 354 AGVLCIALAGLAALPLASVGEYSGASPLCLPYAPPEGOPALGFTVALVMNSCELVY 413
DB 549 AGVLCIALAGLAALPLASVGEYSGASPLCLPYAPPEGOPALGFTVALVMNSCELVY 608
QY 414 AGATIKLYCDLPGRGFEAVWDCAVRYHVAWMLIFADGLLYCPVAFSLFASMLGLFPVTEA 473
DB 609 AGATIKLYCDLPGRGFEAVWDCAVRYHVAWMLIFADGLLYCPVAFSLFASMLGLFPVTEA 668
QY 474 VKSVLLVLPACINPLLYLLFNPHFRDDLRRLPRRGDSGPLAYAAAGLEKSSCDST 533
DB 669 VKSVLLVLPACINPLLYLLFNPHFRDDLRRLPRRGDSGPLAYAAAGLEKSSCDST 728
QY 534 QALVAFSDVDLILEASEGRPPGETTYGPPSVTLISCOOPGAPRLGSGHCVEPCGNHFGN 593
DB 729 QALVAFSDVDLILEASEGRPPGETTYGPPSVTLISCOOPGAPRLGSGHCVEPCGNHFGN 788
QY 594 POPSMDELLLRARGSPAGGSLGGGFGFQSSGLAFASHV 633
DB 789 POPSMDELLLRARGSPAGGSLGGGFGFQSSGLAFASHV 828

RESULT 2
ID LGR5_HUMAN STANDARD; PRT; 907 AA.
AC 075473; 09up75;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
DE (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor
DE 49).
DE GPR49 OR LGR5 OR GPR67.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98308104; PubMed=9642114;
RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T.,
RA Liu Q.;
RT "Identification and cloning of an orphan G protein-coupled receptor of
RT the glycoprotein hormone receptor subfamily."
RL Biochem. Biophys. Res. Commun. 247:266-270(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=placenta;
RX MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a G
RT protein-coupled, seven-transmembrane region."
RL Mol. Endocrinol. 12:1830-1845(1998).
RN [3]
RP FUNCTION: Orphan receptor. It may be an important receptor for
RP signals controlling growth and differentiation of specific
RP embryonic tissues (by similarity).
RN [4]
RP SUBCELLULAR LOCATION: Integral membrane protein.
RN [5]
RP TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal
RN cord, and various region of brain.
RN [6]
RP SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
RN [7]
RP SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF062006; AAC28019.1; -
CC DR EMBL: AF061444; AAC77911.1; -
CC Genew: HGNC:4504; GPR49.

```

```

DR MIM; 606667; 1XUN.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_tyr.
DR Pfam; PF00560; LRR_1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PRO0019; LEURICHPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYR; 8.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
KM Repeat; Leucine-rich repeat.
FT SIGNAL 1 21
FT CHAIN 22 907
FT DOMAIN 1 561
FT TRANSMEM 562 582
FT DOMAIN 583 593
FT TRANSMEM 594 614
FT DOMAIN 615 638
FT TRANSMEM 639 659
FT DOMAIN 660 682
FT TRANSMEM 703 703
FT DOMAIN 722 722
FT TRANSMEM 723 743
FT DOMAIN 744 767
FT TRANSMEM 768 788
FT TRANSMEM 789 802
FT DOMAIN 803 823
FT TRANSMEM 824 907
FT REPEAT 64 88
FT REPEAT 89 112
FT REPEAT 114 136
FT REPEAT 137 160
FT REPEAT 162 184
FT REPEAT 185 208
FT REPEAT 209 232
FT REPEAT 233 256
FT REPEAT 257 279
FT REPEAT 281 303
FT REPEAT 304 327
FT REPEAT 328 350
FT REPEAT 351 375
FT REPEAT 377 396
FT REPEAT 397 420
FT REPEAT 422 444
FT REPEAT 444 468
FT REPEAT 468 492
FT CARBOHYD 564 585
FT CARBOHYD 585 608
FT CARBOHYD 608 631
FT CARBOHYD 631 654
FT CARBOHYD 654 677
FT CARBOHYD 677 700
FT CARBOHYD 700 723
FT CARBOHYD 723 746
FT CARBOHYD 746 769
FT CARBOHYD 769 792
FT CARBOHYD 792 815
FT CARBOHYD 815 838
FT CARBOHYD 838 861
FT CARBOHYD 861 884
FT CARBOHYD 884 907
FT CARBOHYD 907 930
FT CARBOHYD 930 953
FT CARBOHYD 953 976
FT CARBOHYD 976 1000
FT CARBOHYD 1000 1023
FT CARBOHYD 1023 1046
FT CARBOHYD 1046 1069
FT CARBOHYD 1069 1092
FT CARBOHYD 1092 1115
FT CARBOHYD 1115 1138
FT CARBOHYD 1138 1161
FT CARBOHYD 1161 1184
FT CARBOHYD 1184 1207
FT CARBOHYD 1207 1230
FT CARBOHYD 1230 1253
FT CARBOHYD 1253 1276
FT CARBOHYD 1276 1299
FT CARBOHYD 1299 1322
FT CARBOHYD 1322 1345
FT CARBOHYD 1345 1368
FT CARBOHYD 1368 1391
FT CARBOHYD 1391 1414
FT CARBOHYD 1414 1437
FT CARBOHYD 1437 1460
FT CARBOHYD 1460 1483
FT CARBOHYD 1483 1506
FT CARBOHYD 1506 1529
FT CARBOHYD 1529 1552
FT CARBOHYD 1552 1575
FT CARBOHYD 1575 1598
FT CARBOHYD 1598 1621
FT CARBOHYD 1621 1644
FT CARBOHYD 1644 1667
FT CARBOHYD 1667 1690
FT CARBOHYD 1690 1713
FT CARBOHYD 1713 1736
FT CARBOHYD 1736 1759
FT CARBOHYD 1759 1782
FT CARBOHYD 1782 1805
FT CARBOHYD 1805 1828
FT CARBOHYD 1828 1851
FT CARBOHYD 1851 1874
FT CARBOHYD 1874 1897
FT CARBOHYD 1897 1920
FT CARBOHYD 1920 1943
FT CARBOHYD 1943 1966
FT CARBOHYD 1966 1989
FT CARBOHYD 1989 2012
FT CARBOHYD 2012 2035
FT CARBOHYD 2035 2058
FT CARBOHYD 2058 2081
FT CARBOHYD 2081 2104
FT CARBOHYD 2104 2127
FT CARBOHYD 2127 2150
FT CARBOHYD 2150 2173
FT CARBOHYD 2173 2196
FT CARBOHYD 2196 2219
FT CARBOHYD 2219 2242
FT CARBOHYD 2242 2265
FT CARBOHYD 2265 2288
FT CARBOHYD 2288 2311
FT CARBOHYD 2311 2334
FT CARBOHYD 2334 2357
FT CARBOHYD 2357 2380
FT CARBOHYD 2380 2403
FT CARBOHYD 2403 2426
FT CARBOHYD 2426 2449
FT CARBOHYD 2449 2472
FT CARBOHYD 2472 2495
FT CARBOHYD 2495 2518
FT CARBOHYD 2518 2541
FT CARBOHYD 2541 2564
FT CARBOHYD 2564 2587
FT CARBOHYD 2587 2610
FT CARBOHYD 2610 2633
FT CARBOHYD 2633 2656
FT CARBOHYD 2656 2679
FT CARBOHYD 2679 2702
FT CARBOHYD 2702 2725
FT CARBOHYD 2725 2748
FT CARBOHYD 2748 2771
FT CARBOHYD 2771 2794
FT CARBOHYD 2794 2817
FT CARBOHYD 2817 2840
FT CARBOHYD 2840 2863
FT CARBOHYD 2863 2886
FT CARBOHYD 2886 2909
FT CARBOHYD 2909 2932
FT CARBOHYD 2932 2955
FT CARBOHYD 2955 2978
FT CARBOHYD 2978 3001
FT CARBOHYD 3001 3024
FT CARBOHYD 3024 3047
FT CARBOHYD 3047 3070
FT CARBOHYD 3070 3093
FT CARBOHYD 3093 3116
FT CARBOHYD 3116 3139
FT CARBOHYD 3139 3162
FT CARBOHYD 3162 3185
FT CARBOHYD 3185 3208
FT CARBOHYD 3208 3231
FT CARBOHYD 3231 3254
FT CARBOHYD 3254 3277
FT CARBOHYD 3277 3300
FT CARBOHYD 3300 3323
FT CARBOHYD 3323 3346
FT CARBOHYD 3346 3369
FT CARBOHYD 3369 3392
FT CARBOHYD 3392 3415
FT CARBOHYD 3415 3438
FT CARBOHYD 3438 3461
FT CARBOHYD 3461 3484
FT CARBOHYD 3484 3507
FT CARBOHYD 3507 3530
FT CARBOHYD 3530 3553
FT CARBOHYD 3553 3576
FT CARBOHYD 3576 3599
FT CARBOHYD 3599 3622
FT CARBOHYD 3622 3645
FT CARBOHYD 3645 3668
FT CARBOHYD 3668 3691
FT CARBOHYD 3691 3714
FT CARBOHYD 3714 3737
FT CARBOHYD 3737 3760
FT CARBOHYD 3760 3783
FT CARBOHYD 3783 3806
FT CARBOHYD 3806 3829
FT CARBOHYD 3829 3852
FT CARBOHYD 3852 3875
FT CARBOHYD 3875 3898
FT CARBOHYD 3898 3921
FT CARBOHYD 3921 3944
FT CARBOHYD 3944 3967
FT CARBOHYD 3967 3990
FT CARBOHYD 3990 4013
FT CARBOHYD 4013 4036
FT CARBOHYD 4036 4059
FT CARBOHYD 4059 4082
FT CARBOHYD 4082 4105
FT CARBOHYD 4105 4128
FT CARBOHYD 4128 4151
FT CARBOHYD 4151 4174
FT CARBOHYD 4174 4197
FT CARBOHYD 4197 4220
FT CARBOHYD 4220 4243
FT CARBOHYD 4243 4266
FT CARBOHYD 4266 4289
FT CARBOHYD 4289 4312
FT CARBOHYD 4312 4335
FT CARBOHYD 4335 4358
FT CARBOHYD 4358 4381
FT CARBOHYD 4381 4404
FT CARBOHYD 4404 4427
FT CARBOHYD 4427 4450
FT CARBOHYD 4450 4473
FT CARBOHYD 4473 4496
FT CARBOHYD 4496 4519
FT CARBOHYD 4519 4542
FT CARBOHYD 4542 4565
FT CARBOHYD 4565 4588
FT CARBOHYD 4588 4611
FT CARBOHYD 4611 4634
FT CARBOHYD 4634 4657
FT CARBOHYD 4657 4680
FT CARBOHYD 4680 4703
FT CARBOHYD 4703 4726
FT CARBOHYD 4726 4749
FT CARBOHYD 4749 4772
FT CARBOHYD 4772 4795
FT CARBOHYD 4795 4818
FT CARBOHYD 4818 4841
FT CARBOHYD 4841 4864
FT CARBOHYD 4864 4887
FT CARBOHYD 4887 4910
FT CARBOHYD 4910 4933
FT CARBOHYD 4933 4956
FT CARBOHYD 4956 4979
FT CARBOHYD 4979 5002
FT CARBOHYD 5002 5025
FT CARBOHYD 5025 5048
FT CARBOHYD 5048 5071
FT CARBOHYD 5071 5094
FT CARBOHYD 5094 5117
FT CARBOHYD 5117 5140
FT CARBOHYD 5140 5163
FT CARBOHYD 5163 5186
FT CARBOHYD 5186 5209
FT CARBOHYD 5209 5232
FT CARBOHYD 5232 5255
FT CARBOHYD 5255 5278
FT CARBOHYD 5278 5301
FT CARBOHYD 5301 5324
FT CARBOHYD 5324 5347
FT CARBOHYD 5347 5370
FT CARBOHYD 5370 5393
FT CARBOHYD 5393 5416
FT CARBOHYD 5416 5439
FT CARBOHYD 5439 5462
FT CARBOHYD 5462 5485
FT CARBOHYD 5485 5508
FT CARBOHYD 5508 5531
FT CARBOHYD 5531 5554
FT CARBOHYD 5554 5577
FT CARBOHYD 5577 5600
FT CARBOHYD 5600 5623
FT CARBOHYD 5623 5646
FT CARBOHYD 5646 5669
FT CARBOHYD 5669 5692
FT CARBOHYD 5692 5715
FT CARBOHYD 5715 5738
FT CARBOHYD 5738 5761
FT CARBOHYD 5761 5784
FT CARBOHYD 5784 5807
FT CARBOHYD 5807 5830
FT CARBOHYD 5830 5853
FT CARBOHYD 5853 5876
FT CARBOHYD 5876 5899
FT CARBOHYD 5899 5922
FT CARBOHYD 5922 5945
FT CARBOHYD 5945 5968
FT CARBOHYD 5968 5991
FT CARBOHYD 5991 6014
FT CARBOHYD 6014 6037
FT CARBOHYD 6037 6060
FT CARBOHYD 6060 6083
FT CARBOHYD 6083 6106
FT CARBOHYD 6106 6129
FT CARBOHYD 6129 6152
FT CARBOHYD 6152 6175
FT CARBOHYD 6175 6198
FT CARBOHYD 6198 6221
FT CARBOHYD 6221 6244
FT CARBOHYD 6244 6267
FT CARBOHYD 6267 6290
FT CARBOHYD 6290 6313
FT CARBOHYD 6313 6336
FT CARBOHYD 6336 6359
FT CARBOHYD 6359 6382
FT CARBOHYD 6382 6405
FT CARBOHYD 6405 6428
FT CARBOHYD 6428 6451
FT CARBOHYD 6451 6474
FT CARBOHYD 6474 6497
FT CARBOHYD 6497 6520
FT CARBOHYD 6520 6543
FT CARBOHYD 6543 6566
FT CARBOHYD 6566 6589
FT CARBOHYD 6589 6612
FT CARBOHYD 6612 6635
FT CARBOHYD 6635 6658
FT CARBOHYD 6658 6681
FT CARBOHYD 6681 6704
FT CARBOHYD 6704 6727
FT CARBOHYD 6727 6750
FT CARBOHYD 6750 6773
FT CARBOHYD 6773 6796
FT CARBOHYD 6796 6819
FT CARBOHYD 6819 6842
FT CARBOHYD 6842 6865
FT CARBOHYD 6865 6888
FT CARBOHYD 6888 6911
FT CARBOHYD 6911 6934
FT CARBOHYD 6934 6957
FT CARBOHYD 6957 6980
FT CARBOHYD 6980 7003
FT CARBOHYD 7003 7026
FT CARBOHYD 7026 7049
FT CARBOHYD 7049 7072
FT CARBOHYD 7072 7095
FT CARBOHYD 7095 7118
FT CARBOHYD 7118 7141
FT CARBOHYD 7141 7164
FT CARBOHYD 7164 7187
FT CARBOHYD 7187 7210
FT CARBOHYD 7210 7233
FT CARBOHYD 7233 7256
FT CARBOHYD 7256 7279
FT CARBOHYD 7279 7302
FT CARBOHYD 7302 7325
FT CARBOHYD 7325 7348
FT CARBOHYD 7348 7371
FT CARBOHYD 7371 7394
FT CARBOHYD 7394 7417
FT CARBOHYD 7417 7440
FT CARBOHYD 7440 7463
FT CARBOHYD 7463 7486
FT CARBOHYD 7486 7509
FT CARBOHYD 7509 7532
FT CARBOHYD 7532 7555
FT CARBOHYD 7555 7578
FT CARBOHYD 7578 7601
FT CARBOHYD 7601 7624
FT CARBOHYD 7624 7647
FT CARBOHYD 7647 7670
FT CARBOHYD 7670 7693
FT CARBOHYD 7693 7716
FT CARBOHYD 7716 7739
FT CARBOHYD 7739 7762
FT CARBOHYD 7762 7785
FT CARBOHYD 7785 7808
FT CARBOHYD 7808 7831
FT CARBOHYD 7831 7854
FT CARBOHYD 7854 7877
FT CARBOHYD 7877 7900
FT CARBOHYD 7900 7923
FT CARBOHYD 7923 7946
FT CARBOHYD 7946 7969
FT CARBOHYD 7969 7992
FT CARBOHYD 7992 8015
FT CARBOHYD 8015 8038
FT CARBOHYD 8038 8061
FT CARBOHYD 8061 8084
FT CARBOHYD 8084 8107
FT CARBOHYD 8107 8130
FT CARBOHYD 8130 8153
FT CARBOHYD 8153 8176
FT CARBOHYD 8176 8199
FT CARBOHYD 8199 8222
FT CARBOHYD 8222 8245
FT CARBOHYD 8245 8268
FT CARBOHYD 8268 8291
FT CARBOHYD 8291 8314
FT CARBOHYD 8314 8337
FT CARBOHYD 8337 8360
FT CARBOHYD 8360 8383
FT CARBOHYD 8383 8406
FT CARBOHYD 8406 8429
FT CARBOHYD 8429 8452
FT CARBOHYD 8452 8475
FT CARBOHYD 8475 8498
FT CARBOHYD 8498 8521
FT CARBOHYD 8521 8544
FT CARBOHYD 8544 8567
FT CARBOHYD 8567 8590
FT CARBOHYD 8590 8613
FT CARBOHYD 8613 8636
FT CARBOHYD 8636 8659
FT CARBOHYD 8659 8682
FT CARBOHYD 8682 8705
FT CARBOHYD 8705 8728
FT CARBOHYD 8728 8751
FT CARBOHYD 8751 8774
FT CARBOHYD 8774 8797
FT CARBOHYD 8797 8820
FT CARBOHYD 8820 8843
FT CARBOHYD 8843 8866
FT CARBOHYD 8866 8889
FT CARBOHYD 8889 8912
FT CARBOHYD 8912 8935
FT CARBOHYD 8935 8958
FT CARBOHYD 8958 8981
FT CARBOHYD 8981 9004
FT CARBOHYD 9004 9027
FT CARBOHYD 9027 9050
FT CARBOHYD 9050 9073
FT CARBOHYD 9073 9096
FT CARBOHYD 9096 9119
FT CARBOHYD 9119 9142
FT CARBOHYD 9142 9165
FT CARBOHYD 9165 9188
FT CARBOHYD 9188 9211
FT CARBOHYD 9211 9234
FT CARBOHYD 9234 9257
FT CARBOHYD 9257 9280
FT CARBOHYD 9280 9303
FT CARBOHYD 9303 9326
FT CARBOHYD 9326 9349
FT CARBOHYD 9349 9372
FT CARBOHYD 9372 9395
FT CARBOHYD 9395 9418
FT CARBOHYD 9418 9441
FT CARBOHYD 9441 9464
FT CARBOHYD 9464 9487
FT CARBOHYD 9487 9510
FT CARBOHYD 9510 9533
FT CARBOHYD 9533 9556
FT CARBOHYD 9556 9579
FT CARBOHYD 9579 9602
FT CARBOHYD 9602 9625
FT CARBOHYD 9625 9648
FT CARBOHYD 9648 9671
FT CARBOHYD 9671 9694
FT CARBOHYD 9694 9717
FT CARBOHYD 9717 9740
FT CARBOHYD 9740 9763
FT CARBOHYD 9763 9786
FT CARBOHYD 9786 9809
FT CARBOHYD 9809 9832
FT CARBOHYD 9832 9855
FT CARBOHYD 9855 9878
FT CARBOHYD 9878 9901
FT CARBOHYD 9901 9924
FT CARBOHYD 9924 9947
FT CARBOHYD 9947 9970
FT CARBOHYD 9970 9993
FT CARBOHYD 9993 10016
FT CARBOHYD 10016 10039
FT CARBOHYD 10039 10062
FT CARBOHYD 10062 10085
FT CARBOHYD 10085 10108
FT CARBOHYD 10108 10131
FT CARBOHYD 10131 10154
FT CARBOHYD 10154 10177
FT CARBOHYD 10177 10200
FT CARBOHYD 10200 10223
FT CARBOHYD 10223 10246
FT CARBOHYD 10246 10269
FT CARBOHYD 10269 10292
FT CARBOHYD 10292 10315
FT CARBOHYD 10315 10338
FT CARBOHYD 10338 10361
FT CARBOHYD 10361 10384
FT CARBOHYD 10384 10407
FT CARBOHYD 10407 10430
FT CARBOHYD 10430 10453
FT CARBOHYD 10453 10476
FT CARBOHYD 10476 10499
FT CARBOHYD 10499 10522
FT CARBOHYD 10522 10545
FT CARBOHYD 10545 10568
FT CARBOHYD 10568 10591
FT CARBOHYD 10591 10614
FT CARBOHYD 10614 10637
FT CARBOHYD 10637 10660
FT CARBOHYD 10660 10683
FT CARBOHYD 10683 10706
FT CARBOHYD 10706 10729
FT CARBOHYD 10729 10752
FT CARBOHYD 10752 10775
FT CARBOHYD 10775 10798
FT CARBOHYD 10798 10821
FT CARBOHYD 10821 10844
FT CARBOHYD 10844 10867
FT CARBOHYD 10867 10890
FT CARBOHYD 10890 10913
FT CARBOHYD 10913 10936
FT CARBOHYD 10936 10959
FT CARBOHYD 10959 10982
FT CARBOHYD 10982 11005
FT CARBOHYD 11005 11028
FT CARBOHYD 11028 11051
FT CARBOHYD 11051 11074
FT CARBOHYD 11074 11097
FT CARBOHYD 11097 11120
FT CARBOHYD 11120 11143
FT CARBOHYD 11143 11166
FT CARBOHYD 11166 11189
FT CARBOHYD 11189 11212
FT CARBOHYD 11212 11235
FT CARBOHYD 11235 11258
FT CARBOHYD 11258 11281
FT CARBOHYD 11281 11304
FT CARBOHYD 11304 11327
FT CARBOHYD 11327 11350
FT CARBOHYD 11350 11373
FT CARBOHYD 11373 11396
FT CARBOHYD 11396 11419
FT CARBOHYD 11419 11442
FT CARBOHYD 11442 11465
FT CARBOHYD 11465 11488
FT CARBOHYD 11488 11511
FT CARBOHYD 11511 11534
FT CARBOHYD 11534 11557
FT CARBOHYD 11557 11580
FT CARBOHYD 11580 11603
FT CARBOHYD 11603 11626
FT CARBOHYD 11626 11649
FT CARBOHYD 11649 11672
FT CARBOHYD 11672 11695
FT CARBOHYD 11695 11718
FT CARBOHYD 11718 11741
FT CARBOHYD 11741 11764
FT CARBOHYD 11764 11787
FT CARBOHYD 11787 11810
FT CARBOHYD 11810 11833
FT CARBOHYD 11833 11856
FT CARBOHYD 11856 11879
FT CARBOHYD 11879 11902
FT CARBOHYD 11902 11925
FT CARBOHYD 11925 11948
FT CARBOHYD 11948 11971
FT CARBOHYD 11971 11994
FT CARBOHYD 11994 12017
FT CARBOHYD 12017 12040
FT CARBOHYD 12040 12063
FT CARBOHYD 12063 12086
FT CARBOHYD 12086 12109
FT CARBOHYD 12109 12132
FT CARBOHYD 12132 12155
FT CARBOHYD 12155 12178
FT CARBOHYD 12178 12201
FT CARBOHYD 12201 12224
FT CARBOHYD 12224 12247
FT CARBOHYD 12247 12270
FT CARBOHYD 12270 12293
FT CARBOHYD 12293 12316
FT CARBOHYD 12316 12339
FT CARBOHYD 12339 12362
FT CARBOHYD 12362 12385
FT CARBOHYD 12385 12408
FT CARBOHYD 12408 12431
FT CARBOHYD 12431 12454
FT CARBOHYD 12454 12477
FT CARBOHYD 12477 12500
FT CARBOHYD 12500 12523
FT CARBOHYD 12523 12546
FT CARBOHYD 12546 12569
FT CARBOHYD 12569 12592
FT CARBOHYD 12592 12615
FT CARBOHYD 12615 12638
FT CARBOHYD 12638 12661
FT CARBOHYD 12661 12684
FT CARBOHYD 12684 12707
FT CARBOHYD 12707 12730
FT CARBOHYD 12730 12753
FT CARBOHYD 12753 12776
FT CARBOHYD 12776 12799
FT CARBOHYD 12799 12822
FT CARBOHYD 12822 12845
FT CARBOHYD 12845 12868
FT CARBOHYD 12868 12891
FT CARBOHYD 12891 12914
FT CARBOHYD 12914 12937
FT CARBOHYD 12937 12960
FT CARBOHYD 12960 12983
FT CARBOHYD 12983 13006
FT CARBOHYD 13006 13029
FT CARBOHYD 13029 13052
FT CARBOHYD 13052 13075
FT CARBOHYD 13075 13098
FT CARBOHYD 13098 13121
FT CARBOHYD 13121 13144
FT CARBOHYD 13144 13167
FT CARBOHYD 13167 13190
FT CARBOHYD 13190 13213
FT CARBOHYD 13213 13236
FT CARBOHYD 13236 13259
FT CARBOHYD 13259 13282
FT CARBOHYD 13282 13305
FT CARBOHYD 13305 13328
FT CARBOHYD 13328 13351
FT CARBOHYD 13351 13374
FT CARBOHYD 13374 13397
FT CARBOHYD 13397 13420
FT CARBOHYD 13420 13443
FT CARBOHYD 13443 13466
FT CARBOHYD 13466 13489
FT CARBOHYD 13489 13512
FT CARBOHYD 13512 13535
FT CARBOHYD 13535 13558
FT CARBOHYD 13558 13581
FT CARBOHYD 13581 13604
FT CARBOHYD 13604 13627
FT CARBOHYD 13627 13650
FT CARBOHYD 13650 13673
FT CARBOHYD 13673 13696
FT CARBOHYD 13696 13719
FT CARBOHYD 13719 13742
FT CARBOHYD 13742 13765
FT CARBOHYD 13765 13788
FT CARBOHYD 13788 13811
FT CARBOHYD 13811 13834
FT CARBOHYD 13834 13857
FT CARBOHYD 13857 13880
FT CARBOHYD 13880 13903
FT CARBOHYD 13903 13926
FT CARBOHYD 13926 13949
FT CARBOHYD 13949 13972
FT CARBOHYD 13972 13995
FT CARBOHYD 13995 14018
FT CARBOHYD 14018 14041
FT CARBOHYD 14041 14064
FT CARBOHYD 14064 14087
FT CARBOHYD 14087 14110
FT CARBOHYD 14110 14133
FT CARBOHYD 14133 14156
FT CARBOHYD 14156 14179
FT CARBOHYD 14179 14202
FT CARBOHYD 14202 14225
FT CARBOHYD 14225 14248
FT CARBOHYD 14248 14271
FT CARBOHYD 14271 14294
FT CARBOHYD 14294 14317
FT CARBOHYD 14317 14340
FT CARBOHYD 14340 14363
FT CARBOHYD 14363 14386
FT CARBOHYD 14386 14409
FT CARBOHYD 14409 14432
FT CARBOHYD 14432 14455
FT CARBOHYD 14455 14478
FT CARBOHYD 14478 14501
FT CARBOHYD 14501 14524
FT CARBOHYD 14524 14547
FT CARBOHYD 14547 14570
FT CARBOHYD 14570 14593
FT CARBOHYD 14593 14616
FT CARBOHYD 14616 14639
FT CARBOHYD 14639 14662
FT CARBOHYD 14662 14685
FT CARBOHYD 14685 14708
FT CARBOHYD 14708 14731
FT CARBOHYD 14731 14754
FT CARBOHYD 14754 14777
FT CARBOHYD 14777 14800
FT CARBOHYD 14800 14823
FT CARBOHYD 14823 14846
FT CARBOHYD 14846 14869
FT CARBOHYD 14869 14892
FT CARBOHYD 14892 14915
FT CARBOHYD 14915 14938
FT CARBOHYD 14938 14961
FT CARBOHYD 14961 14984
FT CARBOHYD 14984 15007
FT CARBOHYD 15007 15030
FT CARBOHYD 15030 15053
FT CARBOHYD 15053 15076
FT CARBOHYD 15076 15099
FT CARBOHYD 15099 15122
FT CARBOHYD 15122 15145
FT CARBOHYD 15145 15168
FT CARBOHYD 15168 15191
FT CARBOHYD 15191 15214
FT CARBOHYD 15214 15237
FT CARBOHYD 15237 15260
FT CARBOHYD 15260 15283
FT CARBOHYD 15283 15306
FT CARBOHYD 15306 15329
FT CARBOHYD 15329 15352
FT CARBOHYD 15352 15375
FT CARBOHYD 15375 15398
FT CARBOHYD 15398 15421
FT CARBOHYD 15421 15444
FT CARBOHYD 15444 15467
FT CARBOHYD 15467 15490
FT CARBOHYD 15490 15513
FT CARBOHYD 15513 15536
FT CARBOHYD 15536 15559
FT CARBOHYD 15559 15582
FT CARBOHYD 15582 15605
FT CARBOHYD 15605 15628
FT CARBOHYD 15628 15651
FT CARBOHYD 15651 15674
FT CARBOHYD 15674 15697
FT CARBOHYD 15697 15720
FT CARBOHYD 15720 15743
FT CARBOHYD 15743 15766
FT CARBOHYD 15766 15789
FT CARBOHYD 15789 15812
FT CARBOHYD 15812 15835
FT CARBOHYD 15835 15858
FT CARBOHYD 15858 15881
FT CARBOHYD 15881 15904
FT CARBOHYD 15904 15927
FT CARBOHYD 15927 15950
FT CARBOHYD 15950 15973
FT CARBOHYD 15973 15996
FT CARBOHYD 15996 16019
FT CARBOHYD 16019 16042
FT CARBOHYD 16042 16065
FT CARBOHYD 16065 16088
FT CARBOHYD 16088 16111
FT CARBOHYD 16111 16134
FT CARBOHYD 16134 16157
FT CARBOHYD 16157 16180
FT CARBOHYD 16180 16203
FT CARBOHYD 16203 16226
FT CARBOHYD 16226 16249
FT CARBOHYD 16249 16272
FT CARBOHYD 16272 16295
FT CARBOHYD 16295 16318
FT CARBOHYD 16318 16341
FT CARBOHYD 16341 16364
FT CARBOHYD 16364 
```


DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003572; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_Typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_RHODOPN.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_Typ; 5.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; FALSE_NEG.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; signal; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 951
 FT DOMAIN 25 544
 FT TRANSSEM 545 565
 FT DOMAIN 566 575
 FT TRANSSEM 576 596
 FT DOMAIN 597 619
 FT TRANSSEM 620 640
 FT DOMAIN 641 661
 FT TRANSSEM 662 682
 FT DOMAIN 683 703
 FT TRANSSEM 704 724
 FT DOMAIN 725 756
 FT TRANSSEM 757 777
 FT DOMAIN 778 783
 FT TRANSSEM 784 804
 FT DOMAIN 805 951
 FT TRANSSEM 952 979
 FT REPEAT 81 103
 FT REPEAT 104 127
 FT REPEAT 128 151
 FT REPEAT 153 175
 FT REPEAT 176 199
 FT REPEAT 200 223
 FT REPEAT 225 247
 FT REPEAT 248 270
 FT REPEAT 272 294
 FT REPEAT 318 341
 FT REPEAT 342 366
 FT REPEAT 368 387
 FT REPEAT 388 411
 FT REPEAT 413 435
 FT DISULFID 618 693
 FT CARBOHYD 68 68
 FT CARBOHYD 188 188
 FT CARBOHYD 199 199
 FT CARBOHYD 294 294
 FT CARBOHYD 314 314
 FT CARBOHYD 505 505
 SO SEQUENCE 951 AA: 104138 MW; EDD56AC072123461 CRC64;
 Query Match 29.2%; Score 967; DB 1; Length 951;
 Best Local Similarity 39.0%; Pred. No. 1.4e-63;
 Matches 212; Conservative 85; Mismatches 195; Indels 52; Gaps 5;

OY 235 WATVLLSVLNGVLLVFPAGPAPLPVKKVVGAGANTLTGICGLIASVDALTFEQ 294
 DB 546 WFIPLVALLPMLVILVFA SCSSLPASKPIFELISVSNLMQITYGILTFDAVSWGR 604
 OY 295 FSEYGARWETGLGRATGFLAVLGSSEASVLLTLTAQVCSVSVCVAKSPSLGSVA 354
 DB 605 FAELGIMWETSGCKVAGSLAVFSSSEAVFLTLTAVERSFAKDLKHKSHLRFOY 664
 OY 355 GVLGCLALAGLAAMLPLASVGEYGASPLCLPYAPPEGGPALGFTVALVMNSFCPLVA 414
 DB 665 AALLALGAAVAAGCPPLRHGGQYASPLCLPF--PTGFTPSLGFVTLVTLNLSLAFILMA 722
 OY 415 GAVIKYCDLPDPEFAVWDCAMRHVAMILPADGLLYCPVAFISFASMLGFLVTEPAV 474
 DB 723 ILYTLKYLCEKEDLESENSQSSVIRKVAWLIFTNCFECPVAFISFADLYAISPEIM 782
 OY 475 KSVLLVLPPLACNIPILYLPFNPFRDRLRLRPAG----- 512
 DB 783 KSVTLIFPLPACLNPLVLPVFNPKEDMKLKRVRTRKHSVSVSISOGGCEODEY 842
 OY 513 -DSGPLAVAAAGLEKSSCDSTQALVAFSDVLLPABAGRPDELRYGPRSVLISCO 571
 DB 843 YDCGMVSHLQGNLTVCDCESFLTKPKVCKHLI-----KSHSCPVLTAASQ 890
 OY 572 QPGA 575
 DB 891 RPEA 894
 RESULT 5
 LGR4_HUMAN STANDARD; PRT; 951 AA.
 AC QGBXBL; Q9NYD1;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
 DE (G protein-coupled receptor 48).
 GN GPR48 OR LGR4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxID=9606;
 RX MEDLINE=21294803; PubMed=11401528;
 RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
 RT "Molecular characterization of a novel glycoprotein hormone
 RT G-protein-coupled receptor.";
 RT Blochem. Biophys. Res. Commun. 282:757-764(2001).
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
 CC placenta, ovary, testis and adrenal. Expressed also in spinal
 CC cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
 CC and spleen.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF346711; AAK31153.1;
 CC EMBL: AF346709; AAK31153.1; JOINED.
 CC EMBL: AF346710; AAK31153.1; JOINED.
 CC EMBL: AF257182; AAB68989.1; -

DR Genew; HGNC:13299; GPR48.
 DR MIM; 606666; .
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Mterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR; 6.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_typ; 15.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_FL1; FALSE_NEG.
 DR PROSITE: PS00262; G_PROTEIN_RECPT_FL2; 1.
 DR G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
 DR Repeat; Leucine-rich repeat.
 FT SIGNAL 25 951
 FT CHAIN 1 24
 FT DOMAIN 25 544
 FT TRANSMEM 545 565
 FT DOMAIN 566 575
 FT TRANSMEM 576 596
 FT DOMAIN 597 620
 FT TRANSMEM 621 641
 FT DOMAIN 642 661
 FT TRANSMEM 662 682
 FT TRANSMEM 683 703
 FT TRANSMEM 704 724
 FT TRANSMEM 725 756
 FT TRANSMEM 757 777
 FT TRANSMEM 778 783
 FT TRANSMEM 784 804
 FT TRANSMEM 805 951
 FT DOMAIN 55 79
 FT REPEAT 81 103
 FT REPEAT 104 127
 FT REPEAT 128 151
 FT REPEAT 153 175
 FT REPEAT 176 199
 FT REPEAT 201 223
 FT REPEAT 224 247
 FT REPEAT 248 270
 FT REPEAT 272 294
 FT REPEAT 318 341
 FT REPEAT 342 366
 FT REPEAT 368 387
 FT REPEAT 388 411
 FT REPEAT 413 435
 FT REPEAT 433 455
 FT DISULFID 618 693
 FT CARBOHYD 68 199
 FT CARBOHYD 294 294
 FT CARBOHYD 314 314
 FT CARBOHYD 505 505
 FT CONFLICT 292 292
 FT CONFLICT 433 433
 FT CONFLICT 668 668
 FT SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;

Query Match Best Local Similarity 28.9%; Score 957; DB 1; Length 951;
 Matches 211; Conservative 85; Mismatches 191; Indels 60; Gaps 9;

OY 54 KLEKDFESQSLQALDLSMNAIRSHPEAFSTLASHVKKLDTLDTNQLTTLPLAGIGLMH 113
 DB 379 QIKRGTFQGLSLILDLDSRLNLIHEISRATGPTTNLDVSNELTSPFTEGILGILNQ 438
 OY 114 LKLGNLALSGAFSKDSFPKRLILEVPAYVQCCPYGMCASFKAQGWAEADLHLD--- 170

DB 439 LKLVGNFKLAKALAKDFVNIIRSLSPVAYVQCAFWGCDY-----ANLNTEDNSLQDSHV 494
 OY 171 -EESKRRPLGLLQAQENHYDODDELQLEMEDSKPHPSVOCSPTPQPFKCEYLFESWG 229
 DB 495 AOEKGTADAAANVTSLTENEHSQI-----IHCITSTCAFPCCEYLLGSM 540
 OY 230 IRLAVNAIVLLSYCNCLVLLTVFAGPAPLPVKEVYVGAAGANTLTGISCGLLASVDA 289
 DB 541 IRLTWFIPLVAFENLVLITTFASCTSLPSSKLTIFIGLISVNLPMGIYTGITFLDA 599
 OY 290 LTFGQSEYAGRWETGLGCRATGFLAVLGEASVLLTLTAAVOCSSVSVRAGVSPSL 349
 DB 600 VSMGRPAEFQIMWETGSGCKVAGFLAVFESSSAIFLLMTATVRSLSAKDINKRGNH 659
 OY 350 GSVRAVGLGLALAGLAAALPLASVBYGASPLCLPYAPPEGAPALGFTVALVMNSFC 409
 DB 660 KORVVALAFGLGATVAGFPLFRHGEYSASPLCLPF--PTGETPLGFTVTLVLSLA 717
 OY 410 FLVAVAGYIKLYCDLPRGFEAVWDCAMVRHVAWLIFADGLLYCPVAFISFASMLGFLPV 469
 DB 718 FILMAVITYIKLVONLEKEDLSENSQSMKHYAMLIETFCVPAFSPFAPLTAISI 777
 OY 470 TPEAVKSVLLVLPPLACLNPLLYLLENPHRD-----LRLRPRAGDSGPLVAAAGEL 525
 DB 778 SPEIKSVTLIFPLPACLNPLVLYEFNPKFKEDMKLRRYTKKSGSVSISOGGCL 837
 OY 526 EK-----SSCDSTQAIYAFSDYD-----LILEASEGRPPGLTYGFPSTV 566
 DB 838 EODFYDCCGMYSHLQGNLTVCCDCEPFLITKPKVSKHLI-----KSHSCPALA 885
 OY 567 LISQGP 573
 DB 886 VASQCRP 892

RESULT 6
 FSHR_MOUSE STANDARD; PRT; 692 AA.
 ID FSHR_MOUSE
 AC P35378; O9QWV8; Q9DAC2;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (foliitropin receptor).
 GN FSHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV; TISSUE=Testis;
 RA Tena-Sempere M., Manna P.R., Huhtaniemi I.T.;
 RT Molecular cloning of the mouse follicle stimulating hormone receptor complementary deoxyribonucleic acid: functional expression of alternatively spliced variants and receptor inactivation by a C566T transition in exon 7 of the coding sequence.
 RT Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Saito T., Okazaki Y., Nishi K., Kiyosawa H., Kondo S., Yamana K. I., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H., Kuehl P., Lewis S., Matsuo S., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staudt F., Suzuki R., Tomita I., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G., Blake J., Boffelli D., Bojunga N., Carinici P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

[illegible]

1

CC -! SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L22319; AAC37324.1; -
 CC HSSP: P23945; 1XUN.
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR000372; LRR_Nterm.
 CC Pfam: PF00001; 7tm_1; 1.
 CC Pfam: PF00560; LRR_4.
 CC Pfam: PF01462; LRRNT; 1.
 CC SMART: SM00013; LRRNT; 1.
 CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; glycoprotein; signal;
 CC Phosphorylation; Repeat; Leucine-rich repeat.
 KW SIGNAL; 1 17
 FT CHAIN 1 695
 FT DOMAIN 18 366
 FT TRANSSEM 367 387
 FT DOMAIN 388 398
 FT TRANSSEM 399 421
 FT DOMAIN 422 443
 FT TRANSSEM 444 465
 FT DOMAIN 466 485
 FT TRANSSEM 486 508
 FT DOMAIN 509 528
 FT TRANSSEM 529 550
 FT DOMAIN 551 573
 FT TRANSSEM 574 597
 FT DOMAIN 598 608
 FT TRANSSEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 SQ SEQUENCE 695 AA; 78084 MW; 18F9DFEFC046380D CRC64;
 Query Match 12.8%; Score 424.5; DB 1; Length 695;
 Best Local Similarity 26.1%; Pred. No. 9,4e-24;
 Matches 135; Conservative 86; Mismatches 242; Indels 55; Gaps 10;

DB 391 QYKLVPRFIMCNALFADNLGIGYLLLIASVDVHTKTEYHNHVAIDWQIGAGCAAGFTV 450
 QY 317 LGSEASVLLTLAVOCVSVCYRAKSPSLGSVARGVGLAALGLAAALPLASVGE 376
 DB 451 FASELSVTLTALTEHMTTHAMQLECKVQLRHAASIMLVGHIFPAVALFFIFGISS 510
 QY 377 YGASPLCPYAPPEGAPALGFTVALVMNSFCFLVAVAYIKYCDLPRGDF-EAYWDC 435
 DB 511 YKVSICLPM--DISPLSLQLYWMLVNLVLAFAVIGCYTHLYLTVRNPNNTSSSDT 568
 QY 436 AMVRHVAWLIFADGLLCPYAPFLSFASMLGLEPPTPEAVKSVLVLEPLACLNPILYL 495
 DB 569 KIAKRWAMLIETDFLCMAPISEFAISASLKVPLTVSKILLVIFYFINSCANPLTAL 628
 QY 496 FNPFRDRLRLPRAGDSGLVAAAGELKSSCDST 533
 DB 629 FTKNFRDFEILSKFG-----CYEQAQTYRSEHSST 661

RESULT 9
 FSHR_PIG STANDARD; PRT; 695 AA.

AC P49059; 077514;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
 DE receptor).
 GN FSHR
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96011644; Pubmed=7590277;
 RA Remy J.J., Labib-Mansais Y., Verle M., Bozon V., Couture L.,
 RA Payot E., Grebert D., Saleais R.;
 RT "The porcine follicle-stimulating hormone receptor: cDNA cloning, functional
 RT expression and chromosomal localization of the gene.";
 RT Gene 163:257-261(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
 RA la Barbera A.R.;
 RT "Porcine follicle-stimulating hormone receptor.";
 RT Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -! FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -! SUBCELLULAR LOCATION: Integral membrane protein.
 CC -! SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH SUBFAMILY.
 CC -! SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L31966; AA86933.1; -
 CC EMBL: AF025377; AAC24981.1; -
 CC HSSP: P23945; 1XUN.
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR000372; LRR_Nterm.
 CC Pfam: PF00001; 7tm_1; 1.

DR HSSP: P22888; 1LUT.
 DR Interpro: IPR000276; GPCR_Rhodopsn.
 DR Interpro: IPR001611; LRR.
 DR Interpro: IPR000372; LRR_Nterm.
 DR pfam: PF000001; 7tm_1; 1.
 DR pfam: PF000560; LRR_2.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_1; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 27
 FT CHAIN 28 696
 FT DOMAIN 28 358
 FT TRANSMEM 359 386
 FT DOMAIN 387 395
 FT TRANSMEM 396 418
 FT DOMAIN 419 439
 FT TRANSMEM 440 462
 FT DOMAIN 463 482
 FT TRANSMEM 483 505
 FT DOMAIN 506 525
 FT TRANSMEM 526 547
 FT DOMAIN 548 570
 FT TRANSMEM 571 594
 FT DOMAIN 595 605
 FT TRANSMEM 606 626
 FT DOMAIN 627 696
 FT REPEAT 122 147
 FT REPEAT 149 171
 FT REPEAT 172 196
 FT REPEAT 198 220
 FT REPEAT 221 244
 FT DISULFID 439 514
 FT CARBOHYD 99 99
 FT CARBOHYD 174 174
 FT CARBOHYD 195 195
 FT CARBOHYD 291 291
 FT CARBOHYD 299 299
 FT CARBOHYD 313 313
 FT VARSPLIC 317 329
 FT VARSPLIC 330 696
 FT VARSPLIC 317 331
 FT VARSPLIC 332 696
 FT VARSPLIC 317 628
 FT SEQUENCE 696 AA; 78092 MW; 593DEF1C25F982FE CRC64;
 Query Match 12.7%; Score 419; DB 1; Length 696;
 Best Local Similarity 25.28; Pred. No. 2.4e-23;
 Matches 134; Conservative 96; Mismatches 244; Indels 58; Gaps 12;
 QY 26 VERSGSLSPAPASIALAASNTASGKLEXTDFSSQSS-----LOADLSMNA-IR 77
 DB 153 LEICDNLHTTTPAN-ARFGMNNESITLTKNGCFEEQSHAFNGTTLTSLKLNNAHLK 211
 QY 78 SHHPEFTSLHSLVKLDLTNDLTTPLAGLGLMHLKGLMGLAISOAFSKSPFKRLTL 137
 DB 212 KHNDAFRARARPGSIIDISTKLOALPSYGLSIOPLTATSSYLKLPKRREKFTNLDA 271
 QY 138 EYVYAVQCCPY-----GMCASFKASGQWAEFDLHDDSSSKRPLGLARQAEHNY 189
 DB 272 TLVYPSHCCAFRLPTKEQNFSTIRK-----NFSKQCESTARP-----NNETLY 317
 QY 190 -----DQDDELQLEWEDSKPHPSVOCSPTRPGKPCCEYLFEEWGRLAWMAVLVLVLC 244
 DB 318 SAIRFAESELDMQDYDYGFCSF-KTLOCAPEPDAPNCPEDIMGDFLRVLIMLNLIALMG 376
 QY 245 NGVLVLTVPFAGRAPLPVKEVVGAIAGANTLTGISCGLLASVDALTGGFSEYGARWET 304

DB 377 NVTVLEVLTSHYKLTVP-RELMCNLSFADFCMGLYLLLTASVDAQKGGYYNHAIDMOT 435
 QY 305 GLGCRATGFLAULSEASVYLTLTAAVOCVSVCRAVGRKSPSGVAGVGLGCLALAG 364
 DB 436 GNGCSVAGFTVPASELSYTLTTLTTRMHTITVAIGDQDLRLRHAIPIMLGWLFST 495
 QY 365 LAAALPLASVGEYASPLCLPYAPPEGOPALGFYVALVMNMSFCELVAGAYIKLYCDL 424
 DB 496 LIAMPLVGVSSYMKVSIQLPM--DVETTLISQVYITLITLILNVVAFITICACIKITFAY 553
 QY 425 PRGDFEAV-WDCAMVNHVMTLFEADGLYCVAFLSFASMLGLFPYTPRAKSVLLVLP 483
 DB 554 QNPELMATNKTDTIAKMAVLTDTTCAPAPISFFAISAALNKPILITVNSKVLVLFYP 613
 QY 484 LPACLPNPLVLTLPNPHRDLRLRPRAGDSGPLAYAAAGELEKSSDSTQA 535
 DB 614 VNSCANPFLAITTKAFRRDFLL-----LSKSGCKRHOA 648
 RESULT 11
 ID FSHR_SHEEP STANDARD; PRT; 695 AA.
 AC P35379: Q28573; Q28574; Q9TS19;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
 GN FSHR.
 OS Ovis aries (sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R1).
 RC TISSUE=Testis;
 RA MEDLINE=93351750; PubMed=8394255;
 RA Yarney T.A., Sairam M.R., Khan H., Ravindranath N., Payne S., Seidah N.G., "Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor", Mol. Cell. Endocrinol. 93:219-226(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS FSH-R4 AND FSH-R3).
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RA MEDLINE=93176195; PubMed=8439338;
 RA Khan H., Yarney T.A., Sairam M.R.;
 RA "Cloning of alternatively spliced mRNA transcripts coding for variants of ovine testicular follicle receptor lacking the G protein coupling domains", Biochem. Biophys. Res. Commun. 190:888-894(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RA MEDLINE=98031015; PubMed=9364440;
 RA Yarney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W., Sairam M.R.;
 RA "Molecular cloning, structure, and expression of a testicular follicle receptor with selective alteration in the carboxy terminus that affects signaling function.", Mol. Reprod. Dev. 48:458-470(1997).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Ovary;
 RA MEDLINE=20391225; PubMed=10527886;
 RA Babu P.S., Jiang L., Sairam A.M., Touyz R.M., Sairam M.R.;
 RA "Structural features and expression of an alternatively spliced growth factor type I receptor for follicle signaling in the developing ovary.", Mol. Cell Biol. Res. Commun. 2:21-27(1999).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.

[illegible]

*The testicular receptor for follicle stimulating hormone: structure
RT and functional expression of cloned cDNA.*
Mol. Endocrinol. 4:525-530(1990).

SEQUENCE FROM N.A.

MDLINE-92149579; PubMed-1738373;

Heckert L.L., Daley I.J., Griswold M.D.;

*Structural organization of the follicle-stimulating hormone receptor

gene.*

Mol. Endocrinol. 6:70-80(1992).

FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY

OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE

ADENYLATE CYCLASE.

SUBCELLULAR LOCATION: Integral membrane protein.

TISSUE SPECIFICITY: SEROTOLI CELLS AND OVARIAN GRANULOSA CELLS.

SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

FSH/LSH SUPERFAMILY.

SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).

THIS SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

enterprises requires a license agreement (see <http://www.isb.ch/announce/>

or send an email to license@isb.ch).

EMBL; L02842; AAA1175.1.

PIR; A34548; A34548.

PIR; A41729; A41729.

HSSP; P23945; 1XUN.

InterPro: IPR000276; GPCR_Rhodopsn.

InterPro: IPR001611; LRR_Nterm.

InterPro: IPR000372; LRR_Nterm.

PIfam; PF000560; LRR_1.

PIfam; PF01462; LRRNT; 1.

SMART; SM00013; LRRNT; 1.

PROSITE; PS00237; G-PROTEIN_RECPT_F1_1.

PROSITE; PS50262; G-PROTEIN_RECPT_F1_2; 1.

KW G-protein coupled receptor; transmembrane; Glycoprotein; Signal;

Phosphorylation; Repeat; Leucine-rich repeat.

KW POTENTIAL

STGNL 1 17

FT 18 692

FT 18 365

FT 365 386

FT 387 397

FT 420 420

FT 398 442

FT 442 464

FT 464 484

FT 484 507

FT 507 527

FT 527 549

FT 549 572

FT 572 596

FT 596 607

FT 607 629

FT 629 692

FT 692 692

FT 692 692

FT 692 692

FT 692 692

FT 692 692

FT 692 692

FT 692 692

Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

16 SGIPGTHASVERSGG-----SLPAPASIAAALASNTASGKLEXPFSQSL 66

167 Q-----ALDSMNAIRSIHEAP-----S 85

92 SNLPLKLEIRIEKANNLLINPFAONLP-----SLRYLLISNT-----GIKHPAYHKIOSL 144

Db 145 QKVLIDIODININIHVARNSEFGISFESVILNLSKNGIEIRHCACVNGTDELMSDN 204

QY 86 TLHSL-----VRLDLDNQLTTLPLAGIGLMHLKIKGNALSQAFSKDSFPKL 134

Db 205 NLELPNDVPOGASGPVLLDISRTKVHSLPNHGLENLKRIARSTRYLRKLPLDKFVTL 264

QY 135 RILEVPAYOCPCYGMCAAFKASGOWEADLH-----LD 169

Db 265 MEASLTPYSHCAF-----ANLRQISELHPTCKNSILRODIDMTIGORVSLD 316

QY 170 DESSKRPGLGLARQENHYDODLDELQLEMEDSKPHPSVOCSPTPGPFKPEYLEFSWG 229

Db 317 DEPS-----YKGSDDMYNEDYDLN--EYVD-----VTCSPKPDAFNFCEDIMGYNT 363

QY 230 IRLAVAVILVLCNGVLTIVFAGPAPLPVVKFVGATAGANTLGISGLASVDA 289

Db 364 LRVILWFSLIATIGNTVLV-VLTISQYKLTVPRLMCLNLFADLCIGIYLLIASVDI 422

QY 290 LFFGPFSEYGARWETGLGCRATGFLAYLGEASVLLTLAAVQ-----CSVSYCVRAY 343

Db 423 HRKSOYHNYAIDWOTGACCDAGFFYFASBELSVYTLTALTEIRMTTHAQLDC 478

QY 344 GKPSLGSYRAGVGLCLALAGLAALPLASVGEYASPLCPYAPPEGGPAAALGFTVALV 403

Db 479 -KVQLRHASVWLG-WTFARFAALFPIFGISYKVSICLPM--DIDSPLSQLYMAL 534

QY 404 MANSFCELVVAGRYIKIVCDLPRGDF-EAVWDCAVHRVAMLIAPADLLCYPAFLSPAS 462

Db 535 VLVNVAEVVIGCTHYITITVNPPLIVSSSDPTKIARKMATLITDLCAPLSPFAISA 594

QY 463 MGLFPTPEAVKSVLVVLPAPACINPLLYLFNPHRDDLRRLRPAGDSGLAVAA 522

Db 595 SLKPLPLTVSKAKILLVLEYPINSCANPLVLAIFKRNRODFIL-----639

QY 523 GELEKSCSDQVLAVSDVDLLLEASSEAGRPGLFTYFEPVTLISCOGAPRLSESH 582

Db 640 -LSKFGCYEMQAOIYNT-----ETISSA-----TNF-HAKSRHS--SAPVTSY 681

QY 583 CYEPEGNH 590

Db 682 VLAVL-NH 688

RESULT 13
FSHR_HUMAN STANDARD: PRT: 695 AA.

AC P23945;

DT 01-MAR-1992 (rel. 21, Created)

DT 01-JUN-1994 (rel. 29, Last sequence update)

DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle

receptor).

GN FSHR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euteleia; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE-Ovary;

MDLINE-91222171; PubMed-1709010;

RX Mingosha T., Nakamura K., Takakura Y., Igarashi M.;

"Cloning and sequencing of human FSH receptor cDNA."

Biochem. Biophys. Res. Commun. 175:1125-1130(1991).

RT [2]

Query Match

Best Local Similarity

12.6%; Score 417; DB 1; Length 692;

25.0%; Pred. No. 3.3e-23;

RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA MEDLINE=93246012; PubMed=1301382;
 RA Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzesja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 RT its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Alhara T., Nishimori K., Jai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E., Hsueh A.J.;
 RL Submitted (xxx-1992) to the EMBL/Genbank/DBJ databases.
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE-Testis;
 RA MEDLINE=93075197; PubMed=1359889;
 RA Gromoll J., Gudermand T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 RT stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RP SEQUENCE OF 1-51 FROM N.A.
 RA MEDLINE=95011044; PubMed=7926278;
 RA Gromoll J., Dankbar B., Gudermand T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 RT stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 RP 3D-STRUCTURE MODELING OF 49-228.
 RA MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 RT hormone receptors and for the nature of hormone-receptor interactions";
 RL Structure 3:1341-1353(1995).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: SEROTOLI CELLS AND OVARIAN GRAVULOSA CELLS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M65085; AAAS2477.1;
 DR EMBL: S59800; AAB26480.1;
 DR EMBL: M95489; AAAS2478.1;
 DR EMBL: X68044; CA448179.1;
 DR EMBL: S73199; AAB32071.1;
 DR PIR: JN0122; JN0122;
 DR PDB: 1XUN; 15-MAY-97.
 DR Genew; HGNC:3969; FSHR.
 DR MIM: 136435;
 DR Interpro: IPR000276; GPCR_Rhodpsn.
 DR Interpro: IPR001611; LRR.
 DR Interpro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SMO0013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_FL1; 1.

DR PROSITE: PS0262; G-PROTEIN_RECPT_FL2; 1.
 KW G-PROTEIN COUPLED RECEPTOR; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
 KW 3D-structure.
 FT SIGNAL 1 17
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT CARBOHYD 318 318
 FT VARSPLIC 224 285
 FT VARSPLIC 342 695
 FT CONFLICT 13 13
 FT CONFLICT 112 112
 FT CONFLICT 197 198
 FT CONFLICT 295 295
 FT CONFLICT 307 307
 FT CONFLICT 680 680
 FT SEQUENCE 695 AA; 78294 MW; 72388EF1F6DCD5 CRC64;
 Query Match 12.5%; Score 414.5; DB 1; Length 695;
 Best Local Similarity 25.0%; Pred. No. 5.1e-23;
 Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;
 QY 55 LEXDTPSOIS-SIQLNDLSWMAIRSIHPEAF-----STLHSL----- 90
 DB 160 IERNSPVGLSFESVILIMLNKNGIOEIHNCACFNGQLDELNSDNNNLELPNDVFGASC 219
 QY 91 -VKDLITNQLTTPPLAGLIGLIMHLKLNLAQASKSQSFRLILEVPAVYCCPYG 149
 DB 220 PVLIDISRTKRHSIPYGLLENLKKLRASSTYNLKKLPTLERLVALMASLTPSHCCAF- 278
 QY 150 MCASFASGQWEAEDH-----LDDESSRRPLGLARQENHYODDDELQ 198
 DB 279 -----ANMRQISLHPICNKSNIIRQEVDMYOTRGQSSLAEDNESSYRGCDMYT 331
 QY 199 EMDSKRHP--SVQSETPGPFKCEYLFESWGIKRLAWAIVLSYLCGLVLTVFAGG 256
 DB 332 EFDYDLCNEVVDYCSKPKPAFNPCEIDINGYNILKYLMEISILAITGN-IIVILYLT 390
 QY 257 PAPLPVKFVVAAGANLTLTSGGLASVDALTFGGFSEYGARWETGLCRMTFLAY 316
 DB 391 QYKLTVPFRFLMCLNLAADICITIVLLIASVDIHRKSYHNHYAIDMOWGAGCDAAGFFV 450
 QY 317 IGSEASVLLTLTAAYO-----CSVSCVAVYAGKSPSLGVRAGVGCGLALGLAAALP 370
 DB 451 FASELSVTLTLTLERHHTTHAMQDLC-----KYDLRAASVAVWGWT-FAFAALFP 504
 QY 371 LASVEYGAAPLCIPYAPPEGQPAALGFTVALVMAVNSFCGLVAVAGATIKLDCPLPGDF- 429
 DB 505 IFGISSYMKVASICLPM--DIDSPISQLYVMSLVLNVLAVVVICGYIHLYLVRNPNIIV 562

QY 430 EAWDCAMVHVAVMLIFADGLLYCPVALFSPASMLGFPVTPBEAKSVLLVPLPACIN 489
 DB 563 SSSSDPRIAKRAMMLIFDFLCMAPIFAISASLKVPILITYSKAKILLVLPINSCAN 622
 QY 490 PLLYLLFNPHFRDDLRRLPRAGDSGLPAAAGELKSSCDST 533
 DB 623 PFLYALFTKNFRDFFILLSKG-----CYEMOQIYRTETST 661

RESULT 14
 FSHR_MACFA STANDARD; PRT: 695 AA.
 ID FSHR_MACFA
 AC P3212;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
 GN FSHR.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 OX NCBI_Taxid=9541:
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=94071854; PubMed=7504463;
 RA Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;
 RT "Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis."
 RT Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X74454; CAA52463.1; .
 CC PIR: S36452; S36452.
 CC PIR: JN0898; JN0898.
 DR HSP: P23945; LXUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR_Nterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_3.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
 DR G-protein coupled receptor; transmembrane; glycoprotein; signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT STGNL 1 17
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSSEM 367 387
 FT DOMAIN 388 398
 FT TRANSSEM 399 421
 FT DOMAIN 422 443
 FT TRANSSEM 444 465
 FT DOMAIN 466 485

FT TRANSSEM 486 508
 FT DOMAIN 509 528
 FT TRANSSEM 529 550
 FT DOMAIN 551 573
 FT TRANSSEM 574 597
 FT DOMAIN 598 608
 FT TRANSSEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT CARBOHYD 318 318
 SQ SEQUENCE 695 AA; 78343 MW; 0D60A23729B5250 CRC64;

Query Match 12.5%; Score 414.5; DB 1; Length 695;
 Best local Similarity 25.3%; Pred. No. 5,1e-23;
 Matches 134; Conservative 89; Mismatches 228; Indels 79; Gaps 13;

QY 55 LEXDTFSQLS-STQALDLSMNAIRSIHPEAF-----STHSLSL----- 90
 DB 160 IERNSEVGLSFESVILMLKNGIQETHNCAFNCTQDELNLSDNNLELPDVRHGASG 219
 QY 91 -VLDLNDLNTPLAGLGLHLKGNLALSOAQSDFPKRLILEPYAYOCPPYG 149
 DB 220 PYLDISRTIHSLPSGLGNLKLKARSYNNKLPLSEKIALVMEASLTPSHCCAF- 278
 QY 150 MCASFERRASQWENADH-----LDDESSKRLPLGLARQAEHYDDLDLQL 198
 DB 279 -----ANMRROISELHPICNKSILRQEVDMYQTRGQSSIAEDNESSYSGFPMYTA 331
 QY 199 EMEBSKPHR--SYOCSPTPGPFKPCXELFESWGIRLAVAVILLVYLCNGLLTVPAGG 256
 DB 332 EFDYDLQNEVVDYDCSKRPDANFCEDILGYNLRLVLPISLAIIGNIIVLVTL-TTS 390
 QY 257 PAPLPYKVVVAGATAGNTLTGISCGLASVDALTFOQSEYGRWETGICRATGFLAV 316
 DB 391 QYKLTVPREFLCNLAFAFDLCIGYLLIASVDHTKSYHNVAIDWOTGACDAGFFTY 450
 QY 317 LGSEASVLLTLAVQ-----CSVSCVRAVYGSBSLGSVRAGVLCCLALAG 364
 DB 451 FASELSVYTLTAITLERWHITTHAMQLDCKVHR-----HAASVVMGWI-FAF 498
 QY 365 LAALPLASVGEYASPLCLPYAPPEQPALGFTVALVMNSFCPLVAVAGATIKLXCDL 424
 DB 499 AALPFIIGISYMKVSLCLPM--DISPLSOLYVMSLLVNLAVFVIGCYTHIYLV 556
 QY 425 PRGDE-EAWDCAMVHVAVMLIFADGLLYCPVALFSPASMLGFPVTPBEAKSVLLVLP 483
 DB 557 RNPNIYSSSDPRIAKRAMMLIFDFLCMAPIFAISASLKVPILITYSKAKILLVLP 616
 QY 484 LPACNPLLYLLFNPHFRDDLRRLPRAGDSGLPAAAGELKSSCDST 533
 DB 617 INSCANPFLYALFTKNFRDFFILLSKG-----CYEMOQIYRTETST 661

RESULT 15
 TSHR_CANFA STANDARD; PRT: 764 AA.
 ID TSHR_CANFA
 AC P14763;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone receptor).
 GN TSHR.
 OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=90098886; PubMed=2602159;
 RA Parentier M., Libert F., Maenhaut C., Lefort A., Gerard C.,
 RA Perret J., van Sande J., Dumont J.E., Vassart G.,
 RT "Nucleotide sequence of the dog thyrotropin receptor cDNA."
 RL Nucleic Acids Res. 17:10493-10493(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=90084524; PubMed=2556796;
 RA Parentier M., Libert F., Maenhaut C., Lefort A., Gerard C.,
 RA Perret J., van Sande J., Dumont J.E., Vassart G.,
 RT "Molecular cloning of the thyrotropin receptor."
 RL Science 246:1620-1622(1989).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
 CC CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17146; CAA35026.1; -
 DR EMBL: X17147; CAA35027.1; -
 DR EMBL: M29957; AAA30901.1; -
 DR EMBL: M90047; AAA30902.1; -
 DR PIR: S06933; S06933.
 DR PIR: A40077; A40077.
 DR HSSP: P16473; 1XUM.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00560; LRR_1; 1.
 DR Pfam: PF00560; LRR_2; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_FL_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEPTOR_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 764
 FT DOMAIN 21 413
 FT TRANSSEM 414 441
 FT DOMAIN 442 450
 FT TRANSSEM 451 473
 FT DOMAIN 474 494
 FT TRANSSEM 495 517
 FT DOMAIN 518 537
 FT TRANSSEM 538 560
 FT DOMAIN 561 580
 FT TRANSSEM 581 602
 FT DOMAIN 603 625
 FT TRANSSEM 626 649
 FT DOMAIN 650 660
 FT TRANSSEM 661 682
 FT DOMAIN 683 764
 FT REPEAT 51 74
 FT REPEAT 150 174
 FT REPEAT 176 199
 FT REPEAT 201 223
 FT DISULFID 494 569
 BY SIMILARITY.

FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 81 105 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 764 AA; 86483 MW; 49F03B3DBC85512 CRC64;
 Query Match 12.5%; Score 414.5; DB 1; Length 764;
 Best Local Similarity 23.7%; Pred. NO. 5.7e-23;
 Matches 141; Conservative 102; Mismatches 258; Indels 95; Gaps 12;
 QY 33 SLPAHPASIALAASNTASGKLEDFDSQ-----SIQALDLSMNA-IRSHHPAF 84
 DB 166 SIPAN-----AFQGLCNEELTLKLYNNGTSTIQGHAFNKTKIDAYLNKKYLSAIDKDAF 221
 QY 85 STLS-LVKLDLTDNQLTTPLAGIGIMLKGNLALSOAFSKDSFPKRIEVPYAY 143
 DB 222 GGVYSGPTLDDVSTYVALPSKLEHLKELIARNTWTLKRLPLSLHLTRADLSYPS 281
 QY 144 QCCPY-----GMCASP-----FRASG 159
 DB 282 HCCAFKKQKKIRGLLESIMCNESIRSLRQKSVNTLNGPDDQYEXYLGDSHAGYKDNS 341
 QY 160 QW-----BAEDLHLDDSSSKRPLGLARQAEHNDODLDELQLEMEDSKP 205
 DB 342 QFQDTDSNHYVYFEFEQDEDEILFGQELKNPQETLQARDSHYDYV-----CGG 392
 QY 206 HPSVQSPPTGPKPCPYLFESMGIRLAWAIVILSVLCGVLIVYFAGGAPALPVPKF 265
 DB 393 NEDMVCPRKSDENPCGDINGYFELRVWVSVLLALGVFVILVLTSHYKLYVP-RF 451
 QY 266 VCAIAGANTLTISGGLASVDALTFQGFSEYGARWETGLGCRATGFLVLTSEASVYL 325
 DB 452 LMCNLAFDFCMQMYLLILIASVDLYTHSEYNNHAIQWGTGPGCTNGFFVFEASELSVYT 511
 QY 326 LTAAVQCSVSCVAVYAGSPSLGSRVAGVGLCLALAGLAAALPLASVGEYASPLCLP 385
 DB 512 LTVITLERWATIFAMRLDKIRLRHAYVAIMVGWCCFLALLPLVGISSYAKVSICLP 571
 QY 386 YAPDEGPAALGFTVALVMNSFCFLVAVAGAYIKLYCDLPRGDFE-AVMDCAVVRHVAWL 444
 DB 572 M--DDETPLALVILVLLINIVAFITVSCVYKILITVRNPQYNGDMDTKAKMAVL 629
 QY 445 IFADGLLCPVAFELPSAMLGLFPVTPPEAVKSVLVVLPPLQCLNPLLYLLENPHFRDL 504
 DB 630 ITDECMAPLISFYALSALMKNPLITVNSKILVLFYPLNSCANPFLYAITKAFQRDV 689
 QY 505 RLRLPRAGDSGLPAAVAAAG-----ELEKSSCDSTQALVAFSDVDLLEAS 549
 DB 690 FILLSKFGICKRQAGAVRGQVRSPKNSAGIQIQKVTYRDKRQSLPMWQDEYELLENS 745

Search completed: November 8, 2002, 19:34:53
 Job time : 14.2476 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:55 : Search time 16.1833 Seconds
(without alignments)
3760.229 Million cell updates/sec

Title: US-09-851-595-5
Perfect score: 3307
Sequence: 1 NHTHESMYACRYRSGIPG.....GGLSGGGGPGPSGLAFASHV 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	38.6	907	2	orphan G protein-c
2	1205	36.4	907	2	G protein-coupled
3	429	13.0	694	2	folliotropin recept
4	428.5	13.0	694	2	folliotropin recept
5	424.5	12.8	695	2	follicle stimulat
6	419	12.7	696	2	lutropin-choriogon
7	418.5	12.7	695	2	folliotropin recept
8	417	12.6	692	2	folliotropin recept
9	414.5	12.5	695	1	folliotropin recept
10	414.5	12.5	695	1	folliotropin recept
11	414.5	12.5	764	2	thyroid stimulat
12	414.5	12.5	764	2	thyroid stimulat
13	412.5	12.5	700	2	lutropin receptor
14	408	12.3	764	2	lutropin receptor
15	407.5	12.3	700	2	lutropin receptor
16	407.5	12.3	700	2	lutropin receptor
17	403.5	12.2	764	1	thyroid stimulat
18	400	12.1	814	2	thyroid stimulat
19	397.5	12.0	793	2	thyroid stimulat
20	384.5	11.6	696	2	thyroid stimulat
21	381.5	11.5	764	2	thyroid stimulat
22	381	11.5	699	1	thyroid stimulat
23	366	11.1	889	2	thyroid stimulat
24	308.5	9.3	925	2	thyroid stimulat
25	299	9.0	1115	2	thyroid stimulat
26	149.5	4.5	536	2	lysine carboxypept
27	137.5	4.2	355	1	iodopsin homolog -
28	136.5	4.1	653	2	hypothetical prote
29	131.5	4.0	1784	2	hypothetical prote

30	131	4.0	350	2	13848	Mel-1a melatonin r
31	130	3.9	440	2	JC5520	serotonin receptor
32	130	3.9	440	2	A45121	alpha-1b adrenergi
33	130	3.9	864	2	T08575	protein kinase hom
34	129.5	3.9	367	2	T08349	interferon-inducib
35	129.5	3.9	658	2	B04664	probable receptor-
36	129	3.9	605	2	A41915	insulin-like growt
37	128	3.9	382	2	E84527	hypothetical prote
38	127.5	3.9	382	2	E84527	galanin receptor 2
39	127.5	3.9	387	2	JC5949	silt protein 2 pre
40	127.5	3.9	1469	2	B36665	silt protein 1 pre
41	127.5	3.9	1480	2	A36665	opsin, green-sensi
42	126.5	3.8	349	2	A45229	opsin, green-sensi
43	126.5	3.8	349	2	B45229	opsin, green-sensi
44	125.5	3.8	404	2	JC5784	adrenomedullin rec
45	125	3.8	603	2	JC6128	insulin-like growt

ALIGNMENTS

RESULT 1

JE0176
orphan G protein-coupled receptor precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
C:Accession: JE0176
R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
Biochem. Biophys. Res. Commun. 247, 266-270, 1998
A>Title: Identification and cloning of an orphan G protein-coupled receptor of the gl
A:Reference number: JE0176; MUID:98308104; PMID:9642114
A:Accession: JE0176
A:Molecule type: mRNA
A:Residues: 1-907 <MCD>
A:Cross-references: GB:AF062006; NID:q3366801; PID:AA28019.1; PID:q3366802
A:Comment: This protein is a receptor for a novel class of glycoprotein ligands.
C:Genetics:
A:Gene: HG38
A:Map position: 12q22-23
F:1-21/Domain: signal sequence #status predicted <SIG>
F:562-583/Domain: transmembrane #status predicted <TM1>
F:594-616/Domain: transmembrane #status predicted <TM2>
F:639-660/Domain: transmembrane #status predicted <TM3>
F:681-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 38.6% Score 1275; DB 2; Length 907;
Best local Similarity 46.9% Pred. No. 1.2e-91;
Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;

QY	54	KLEKDPFQSLQSLALDLSNNATRSIHPEAFSLHSLVLDLITDNLITPLAGGLGMH	113
DB	388	EIKVDFQQLSLRSLNLAHMKTAIIHPNAPSLPDIKIDLSNLSPPITGLHGLTH	447
QY	114	LKIKGLAISQAFSKDSFPLRLILEVPIYAOCCPYGMCASFRAASGM-----EAEEDLH	167
DB	448	LKLTGHNALQSLISENPELAKVIEHPYAYOCCAFCEYAKRISIMQWNGDNSSMDLH	507
QY	168	LDDEESKRLGLLROAEHNYDDQLEDEQLDEM-EDSKRHPVSQCSPTPGPKCEYLFE	226
DB	508	-----KKDAGMFOADE---RDLEDFLDEEDIKALHSQCSFSPFPKCEHLD	556
QY	227	SWGIRLAWAIVLISVLCNGLVLLTFVAGGAPALPPYKFFVGAITAGNTLTGSGILLAS	286
DB	557	GWILIRIGWIRIIVLALQNLALVTSVFR-SPLVISPILKILIGYIAAVNMILTGVSSAVLAG	615
QY	287	VDALTFQGFSEYARMTETGCRATGFLAVLGSASVLLITLAAVQCSVSVCAVRAKGS	346
DB	616	VDALTFGSEFARHAGWENGSGCHVIGLISIFASSESVFLLITLAALEGFVSKYSANFEIK	675
QY	347	PSLGSVAGVGLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGGPAALGFTVAIVMMN	406

Sat Nov 9 13:04:29 2002

us-09-851-595-5.rpr

RESULT 4

JC2237

Follicotropin receptor, testis - horse

N:Alternate names: eFSHR (domestic horse)

C:Species: Equus caballus (domestic horse)

C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 13-Aug-1999

C:Accession: J02237; J02370

C:Robert. P.; Amstels, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bidart

R:Robert. P.; Amstels, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bidart

Biochem. Biophys. Res. Commun. 201, 201-207, 1994

A:Title: Cloning and sequencing of the equine testicular follicotropin receptor.

A:Reference number: J02237; MUID:94256960; PMID:8198575

A:Accession: J02237

A:Molecule type: mRNA

A:Residues: 1-694 <R0B>

A:Cross-references: GB:570150; NID:9546896; PID:AMB30854.1; PID:9546897

A:Experimental source: testis

A:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C:Keywords: glycoprotein; hormone receptor; transmembrane protein <LR1>

F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

F:170-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>

F:366-386/Domain: transmembrane #status predicted <TM1>

F:398-420/Domain: transmembrane #status predicted <TM4>

F:443-464/Domain: transmembrane #status predicted <TM5>

F:485-507/Domain: transmembrane #status predicted <TM6>

F:528-549/Domain: transmembrane #status predicted <TM7>

F:573-596/Domain: transmembrane #status predicted <TM8>

F:608-629/Domain: transmembrane #status predicted <TM9>

F:191,199,266,293/Binding site: cardonylate (asn) (covalent) #status predicted

Query Match 15.0%; Score 428.5; DB 2; Length 694;

Best local Similarity 25.2%; Pred. No. 1.8e-25; Indels 95; Gaps 14;

Matches 149; Conservative 94; Mismatches 253; STLSL----- 55

16 SGIGSTHVSVERSGSLPAHPA-----SLAASANT-----TASGKL----- 151
12 SNLPRLHEIRKANNLYTDHDFONLPLOYLLISNGIKHLPVNRHOSLOKVLDDI 151
56 -----EKDFESOLS-SLQALDLSWNAIRSIHPEAF-----STLSL----- 90
152 QDNITHVENRSFESFESTILRLSKNGIOELHNCATGQLODELINSYNNLELPPN 211
91 -----VKLDLTDNOLTTPLAGLGLMHLKGNLALSOAFSDSPKRLILEVPY 141
212 DVFQASGVILLDISGTRHSLPNYGLNKLRLKRASTYNNLKLPSLEKVALMEANLTY 271
142 AYCCCPYGMCAFSSKASG-----QWEADLHLDDESSKRLGLLAR 183
272 PSRHCACF-----ANRRQTSDELQTCNKSILROEVDMDQARGEVSIAEDDESSYKRGDMM 328
184 QAEHHYQDDDELQEDSKPHPSVQCSPTGPPKPCYLFESMGIRLAWAIVLLSVL 243
329 YSEEDYD-----LCNEVYD-----VTCSPKPDAPNCPEDINGYDLRLIWFISILAIT 377
244 CNGVLVTFVAGGAPLPYKVVGAJAGANTLTGSCGLASVDALTGCGSESGARME 303
378 GN-IIVLVILTSOYKRLVPRFLMCLNAPADLCIGYLLIASVDHITFSOHNTAIDM 436
304 TGIGCATGTFLAVLGSSESVLLTLAIVQCSVSVCAVRAVGSFSLGSVAVGLCLALA 363
437 TGAGCDAAEFVTFASSELVYTLTILTRHMTTHAMOLECKVQLRHAASVMVQWTER 496
364 GLAAALPLASVGYGASPLCLPYAPPEGAPALGFTVALVMMNSFCFLVAGAVIKLYCD 423
497 FAVALLPFIESTYMKVSLCPM--DIDSPLOLYVMSLVLNLAFAVIVGCGYIHYTLT 554
424 LPRGDF-EAVVDCAMVRAVWILFADGLLYCPVALSEFASMLGLFPVTPPAVSKVLYVL 482
555 VRNPIVSSSDTKIARMAILITFDPLCMAPISFAISASLKVPLITVSKSLVLYVF 614

RESULT 5

145896

Follicle stimulating hormone receptor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jan-2000

C:Accession: I45896

C:Robert. A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.

Mol. Reprod. Dev. 39, 127-135, 1994

A:Title: Structure of the bovine follicle-stimulating hormone receptor complementary

A:Reference number: I45896; MUID:95127199; PMID:7826612

A:Accession: I45896

A:Molecule type: mRNA

A:Residues: 1-695 <H0U>

A:Cross-references: GB:L22319; NID:9404671; PID:AG37324.1; PID:9404672

A:Experimental source: testis

A:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C:Keywords: glycoprotein; hormone receptor; transmembrane protein <LR1>

F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

F:170-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>

F:366-386/Domain: transmembrane #status predicted <TM1>

F:398-420/Domain: transmembrane #status predicted <TM4>

F:443-464/Domain: transmembrane #status predicted <TM5>

F:485-507/Domain: transmembrane #status predicted <TM6>

F:528-549/Domain: transmembrane #status predicted <TM7>

F:573-596/Domain: transmembrane #status predicted <TM8>

F:608-629/Domain: transmembrane #status predicted <TM9>

Query Match 12.8%; Score 424.5; DB 2; Length 695;

Best local Similarity 26.1%; Pred. No. 3.7e-25; Indels 55; Gaps 10;

Matches 135; Conservative 86; Mismatches 242; STLSL----- 90

55 LKDFESOLS-SLQALDLSWNAIRSIHPEAF-----STLSL----- 151
160 VERNSEFGLFSEEMTWLWSKNGIOELHNCATGQLODELINSDNSNEELNDVFOGASG 219
91 -VKLDLTDNOLTTPLAGLGLMHLKGNLALSOAFSDSPKRLILEVPYAYQCCPYG 149
220 PYLIDISTRIKRLSPYGLNKLRLKRASTYNNLKLPSLEKVALMEANLTY 278
150 MCASFASGOMEADLH-----LDDESSKRLGLLARQAEHHYQDDDELQ 198
279 -----ANRRQTSDELQTCNKSILROEVDMDQARGEVSIAEDDESSYKRGDMM 328
199 EMDSKPHR--SVQCSPTGPPKPCYLFESMGIRLAWAIVLLSVLNCGLVILYFAAG 256
332 EFDYDCNEVYVQTSPEEDAFNCPEDIMODILRLIWFISILAITGNLIVL-ILITS 390
257 PAPPYKVVGAJAGANTLTGSCGLASVDALTGCGSESGARME 303
391 QYKLTVPFLMCLNAPADLCIGYLLIASVDHITFSOHNTAIDM 436
317 LGSEASVLLTLAIVQCSVSVCAVRAVGSFSLGSVAVGLCLALA 363
451 FASELSVTLTATILEKWHITTHAMOLECKVQLRHAASIMLVGMIFARVALFPIFGISS 510
377 YGASPLCLPYAPPEGAPALGFTVALVMMNSFCFLVAGAVIKLYCDLPRGDF-EAVVDC 435
511 YMKVSLCPM--DIDSPLOLYVMSLVLNLAFAVIVGCGYIHYTLVNNPNTSSSDT 568
436 ANVRHVAWILFADGLLYCPVALSEFASMLGLFPVTPPAVSKVLYVLPRACINPLLYL 495
569 KIARMAILITFDPLCMAPISFAISASLKVPLITVSKSLVLYVF 614
496 FNFHFRDLRLRPRAGDSGLPAAAGLEKSSCDST 533
629 FTKNFRDFFILSKFG-----CYEQVQATYRSERST 661

RESULT 6

A11344

Intropin-choriogonadotropin receptor precursor (splice form A) - pig

N:Alternate names: Intropin (hormone-choriogonadotropin receptor)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 21-Jan-2000

C/Accession: A41344
 R/Locus: H. Mistrali, M. Atger, M. Salese, R. Vu Hai-Luu Thi, M.T., Jolivet, A.
 Science 245, 525-528, 1989
 A/Title: Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking transmembrane domain
 A/Accession: A41344
 A/Reference number: A41344
 A/Molecule type: mRNA
 A/Residues: 1-696 <LOC>
 A/Cross-references: GB:M29525; NID:9164528; PID:AAA31062.1; PID:9164529
 C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane protein
 F:1-27/Domain: signal sequence; status predicted <SIG>
 F:28-696/Product: lutropin-choriogonadotropin hormone receptor #status predicted <M>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 12.7% Score 419; DB 2; Length 696;
 Best Local Similarity 25.2%; Pred. No. 1e-24;
 Matches 134; Conservative 96; Mismatches 244; Indels 58; Gaps 12;

26 VERSGSLPAPASIALAALAAWNTASGKLEDTESOLSS-----LQALDLSMNA-IR 77
 153 LEIDNMHTTTPAN-AFOGMNNESTTLKLYGFEETIOHAFNGTTLISLEKENMAHK 211
 -78 SHPEAFSTLSLVKLDLNDQTLTLAGLGLMLKGLNLALSOAFSKDSFPKLRIL 137
 212 KMHDAFRGARPSTLIDISSTKLQALPSYGLSITQTLTASYSYKLPSEKFTNLDA 271
 138 EYVYAYQCCPY-----GMCASFKAAGQWEADLHDESSKRPGLILARQAEHNY 189
 272 TLTYSHCCAFKRNLPKBNFSFIEK-----NFKOCESTARRP-----NNETLY 317
 190 -----DQDLDELQEMEDSKPHPSVOCSPPTGPKCEYLFEESGIRLAVMAVILSYVC 244
 318 SAIRPESLSMDYDYGCPSP-KTLCQAFEPDAFNCEIDIMGYDLKRVILMLINLALAG 376
 245 NGVLLTVFAGGPAAPLPVKEVVGATAGANTLTGISCGLASVDALTGQSEVARET 304
 377 NVYLLTVFVLSHYKLYP-RFLMCLNLSFADFCGGLYLLIASVDQVHGTQYHNAIDMT 435
 305 GICGRATGLAVLGSEASVLLTLTAAVQCSVSCVRAVGSFSGSVRAVGLCLAG 364
 436 GNGCSVAGFTYVASELSYTLTTLERWHTTITTAIDLDKRLRRAIPMLGCMLEST 495
 365 LAALPLASVGEYASPLCPAPAPGEPALGFTVALVMNNSCFVLVAGVATIKLYCDL 424
 496 LIMPDLVGVSSYMKVSICLPM-DVETTLISOVYLLTLLNVAVFTIICACIKITRAV 553
 425 PGDEEAV-WDCAMVRHVAWLIADGLLYCPVAFSLFASMLGLPVPVAVSVLLVLP 483
 554 ONPELMATNKTQIAKMAVILFTDFCMAPISFPAISALAKPLITVNSKVLVLEYP 613
 484 LPACINPLLYLTFNPHRDRLRLRPAGDSGLAVAAAGELKSSCDSTQA 535
 614 VNSCANPLVATLFTKAFRRDFELL-----LSKGCCKHQA 648

RESULT 7
 JCI493

follicle stimulating hormone receptor - sheep
 N/Alternate names: follicle stimulating hormone receptor
 C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C/Date: 03-Feb-1994 #sequence-revision 03-Feb-1994 #text-change 21-Jan-2000
 C/Accession: JCI493; 147080
 R/Khan, H.; Yarney, T.A.; Saitam, M.R.
 A/Title: Cloning of alternatively spliced mRNA transcripts coding for variants of ovine follicle stimulating hormone receptor
 A/Reference number: JCI493; MUID:93176195; PMID:8439338
 A/Accession: JCI493
 A/Molecule type: mRNA
 A/Residues: 1-695 <KHA>
 A/Cross-references: GB:M29525; NID:9164528; PID:AAA31062.1; PID:9164529
 C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane protein
 F:1-27/Domain: signal sequence; status predicted <SIG>
 F:28-696/Product: lutropin-choriogonadotropin hormone receptor #status predicted <M>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

A/Title: Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor
 A/Reference number: 147080; MUID:93351750; PMID:8394255
 A/Accession: 147080
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-695 <VAR>
 A/Cross-references: GB:L07302; NID:9165884; PID:AAA31525.1; PID:9165885
 C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-27/Domain: signal sequence; status predicted <SIG>
 F:28-696/Product: lutropin-choriogonadotropin hormone receptor #status predicted <M>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 12.7% Score 418.5; DB 2; Length 695;
 Best Local Similarity 26.3%; Pred. No. 1e-24;
 Matches 139; Conservative 80; Mismatches 225; Indels 85; Gaps 13;

55 LKXDFEOLSS-SLOALDSMAIRSHPEAF-----STLHSL----- 90
 160 VERNSPMGLSPESMITVMSKNGIOEHNCAPNGTQDELNLSDNSMLELPMDFOGASG 219
 91 -VKLDLNDQTLTLAGLGLMLKGLNLALSOAFSKDSFPKLRILEVYAYQCCPYG 149
 220 EVLIDISRTIRSLPSYGLLENLKLRAKSTYHLKLPSEKFTVLEASLTYPSCAF- 278
 150 MCASFKAAGQWEADLH-----LDESSKRPGLILARQAEHNYDQDLDELQ 198
 279 -----ANMRQTSIDLHCKNSILROEVDQMTQARQISLADEDEPXYAGCFMMS 331
 199 EMDSD--KPHPSVOCSPPTGPKCEYLFEESGIRLAVMAVILSYCNGVLLTFVAG 256
 332 EFDYDLCSEVVDYDYGCPSP-KTLCQAFEPDAFNCEIDIMGYDLKRVILMLINLALAG 376
 257 PAPLPVKEVVGATAGANTLTGISCGLASVDALTGQSEVARET 304
 391 QYKLVPRFLKCNLAFADLCIGYLLIASVDQVHGTQYHNAIDMT 435
 317 LGSEASVLLTLTAAVQ-----CSVYSC--VRAIKSPSGSVRAVGLCLAGLAA 367
 451 FASELSYTLTALTEERWHTTITTAIDLDKRLRRAIPMLGCMLEST 495
 368 ALPLASVGEYASPLCPAPAPGEPALGFTVALVMNNSCFVLVAGVATIKLYCDL 424
 502 LFPFGLISSYMKVSICLPM-DIDSPSLQVYMSLVNLVAFVIGCTHYTLVVRNP 559
 428 DF-EAVWDCAMVRHVAWLIADGLLYCPVAFSLFASMLGLPVPVAVSVLLVLP 483
 560 NITSSSDTKIKKMAVILFTDFCMAPISFPAISALAKPLITVNSKVLVLEYP 613
 487 CLNPLLYLTFNPHRDRLRLRPAGDSGLAVAAAGELKSSCDSTQA 535
 620 CANPLVATLFTKAFRRDFELL-----LSKGCYEVQA 651

RESULT 8
 A34548

follicle stimulating hormone precursor - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 22-Jan-1993 #sequence-revision 22-Jan-1993 #text-change 13-Aug-1999
 C/Accession: A34548; A41729
 R/Sprengel, R.; Braun, T.; Nikolics, K.; Segaloff, D.L.; Seeburg, P.H.
 A/Title: The testicular receptor for follicle stimulating hormone: structure and function
 A/Reference number: A4548; MUID:91125358; PMID:2126341
 A/Accession: A34548
 A/Molecule type: mRNA
 A/Residues: 1-692 <SPR>
 A/Cross-references: GB:U02842; NID:9204183; PID:AAA41175.1; PID:9204184
 C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-27/Domain: signal sequence; status predicted <SIG>
 F:28-696/Product: lutropin-choriogonadotropin hormone receptor #status predicted <M>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

A:Accession: A41729
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-692 <HEC>
 A:Cross-references: GB:S81198; NID:9245344; PTDN:AA82145.1; PID:9245345
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:81117, NCBIN:81121, NC
 Mol. Endocrinol. 9, 159-170, 1995
 A:Title: Identification of the sites of N-linked glycosylation on the follicle-stimulat
 A:Reference number: A57562; PMID:95295729; PMID:776966
 A:Contents: annotation: glycosylation sites
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follicle
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
 F:1-15/Domain: signal sequence #status predicted <MAV>
 F:16-692/Product: follicle-stimulating hormone binding #status predicted <EHB>
 F:16-366/Domain: extracellular hormone binding repeat homology <LRR1>
 F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F:367-387/Domain: transmembrane #status predicted <TM1>
 F:398-421/Domain: transmembrane #status predicted <TM2>
 F:443-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191,199,293/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:554/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 12.68; Score 417; DB 2; Length 692;

Best local similarity 25.08; Pred. No. 1,4e-24; Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

16 SCIPGSTASVRSGL-----SLPAHPSALAAASNTASGLKEDTFSQSL 66
 92 SNIPKHEIRIEKANLLYINPEAQNLP-----SLRYLLISNT-----GIKLPVAVKIOSL 144
 67 0-----ADLSMAIARSHPEAF-----S 85
 145 QKVLIDIONINHIYARNSEFMGLSESVIWLKNGIEIHNCAFNCTQDELINSDNN 204
 86 TLHSL-----VKLDINDNLTPLAGLGLMHLKGLNLGLSQAFAFKDPSFKL 134
 205 NLEELPNDVFOGASPVYIIDISTKTHSLPNHGLEKTLKARSTYRLKLPNDKFTVL 264
 135 RILEVPAYOCPCYGCASFKAAGWEADLH-----LD 169
 265 MEASLTPHSCAF-----ANLKRQISELHPICNKSILRODIDDMDQIDQVSLID 316
 170 DESSSKRPLGLIARQENHYDQDELQLEMDSKPHSPVQCSPTGPFKPCBYLFESMG 229
 317 DEPS-----YGGKSDMYNEDYDCN-----EVVD-----VTCSPKPDAPNCPDIDINGYNI 363
 230 IRLAWAIVLVSVCNGLVLTVPAGSPADPLPYKPFVVGALNAGANTLTGSCGLASVDA 289
 364 LRLVIMFSLIATIGNFTVLV-VLTTSQYKLTVPRLMCLNARLDCIGIYLLIASVDI 422
 290 LTRGPFSEYGARWETGICRATGFLAVLSEAVLLTLAAVQ-----CSVSYCYRAY 343
 423 HKRSQYHNHNDIMQAGCDAAGFTVYASSELVYTLATILEWHTITTHAMQEC----- 478
 344 GKSPISGVRACVIGCLAGLAAALPLASVGEASGASICLPVAPREGOALDFTVALV 403
 479 -KQJLRHNASVAVLG-WTFAPFAALFPIFGISSYMKVSIPLPM--DIDSPLSQLYVALL 534

404 MNNSPGLVAVAGYIKLYCDLPRGDF-BAVWDCAMRVHVALIFADGLIYCPVAFLEFAS 462
 535 VLVAVFVVIICGTYTHIYLVNPTLVSSSDTKIAKRMATLIDFICMAPISEFAISA 594
 463 MLGLFPTPEAVKSVLLVLPACINPLLYLLFNDHPEDDLRRRLPRAGDSGLPAAVA 522
 595 SLKVPITVSKAKILLVLEFVPINSCANPFLVATFTKPNRDFIL----- 639
 523 GELEKSCDSTQALVAFSDVLDLLEASSEGRPPGLETFGPFSVTLISCOQPGAPRLEGS 582
 640 --LSKFGCYEMQAOIYRT-----ETSSA-----THNF-HARKSHCS--SAPRYTNSY 681
 583 CVEPEGNH 590
 682 VLVPL-NH 688

RESULT 9

ORHUF
 follicle-stimulating hormone precursor - human
 N:Alternative names: follicle stimulating hormone receptor (FSHR)
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence revision 06-Sep-1996 #text change 22-Jun-1999
 C:Accession: 157661; 156448; PC1147; S30560; 157672; JN0122
 R:Gromoll, J.; Dankbar, B.; Gudermann, T.
 Mol. Cell. Endocrinol. 102, 93-102, 1994
 A:Title: Characterization of the 5' flanking region of the human follicle-stimulating
 A:Reference number: 157661; PMID:95011044; PMID:7926278
 A:Accession: 157661
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-51 <GRO>
 A:Cross-references: GB:S73199; NID:9685036; PTDN:AA832071.1; PID:9685037
 R:Gromoll, J.; Ried, T.; Holtgreve-Grez, H.; Nieschlag, E.; Gudermann, T.
 J. Mol. Endocrinol. 12, 265-271, 1994
 A:Title: Localization of the human FSH receptor to chromosome 2 p21 using a genomic P
 A:Reference number: 156448; PMID:9500244; PMID:7916967
 A:Accession: 156448
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 286-695 <GR2>
 A:Cross-references: GB:S73526; NID:9688069; PTDN:AA832225.1; PID:9688070
 R:Gromoll, J.; Gudermann, T.; Nieschlag, E.
 Biochem. Biophys. Res. Commun. 188, 1077-1083, 1992
 A:Title: Molecular cloning of a truncated isoform of the human follicle stimulating h
 A:Reference number: PC1147; PMID:93075197; PMID:1359889
 A:Accession: PC1147
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-223,286-294, 'P', 296-342 <GR3>
 A:Cross-references: EMBL:968044; NID:931473; PTDN:CAA48179.1; PID:931474
 R:Kellon, C.A.; Cheng, S.V.; Nugent, N.P.; Schweickhardt, R.L.; Rosenthal, J.L.; Over
 Mol. Cell. Endocrinol. 89, 141-151, 1992
 A:Title: The cloning of the human follicle stimulating hormone receptor and its expe
 A:Reference number: 157672; PMID:93246012; PMID:1301382
 A:Accession: 157672
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-679, 'N', 681-695 <KEL>
 A:Cross-references: GB:S59900; NID:9300072; PTDN:AA826480.1; PID:9300073
 R:Minagishi, T.; Nakamura, K.; Takakura, Y.; Itoh, Y.; Igarashi, M.
 Biochem. Biophys. Res. Commun. 175, 1125-1130, 1991
 A:Title: Cloning and sequencing of human FSH receptor cDNA.
 A:Reference number: JN0122; PMID:91222171; PMID:11709010

A:Accession: JN0122
 A:Molecule type: mRNA
 A:Residues: 1-111, 'T', 113-196, 'AV', 199-306, 'A', 308-695 <MIN>
 A:Cross-references: EMBL:565085; NID:9182770; PID:AAA52477.1; PID:9182771
 C:Genetics:
 A:Gene: GDB:FSHR
 A:Cross-references: GDB:127510; OMIM:136435
 A:Map position: 2p21-2p16
 A:Note: The exact position of the intron cannot be determined from the experimental data
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follitropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone recep
 F:1-695/Product: follitropin receptor precursor, long splice form #status predicted <spi
 F:1-223,286-695/Product: follitropin receptor precursor, short splice form #status predi
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-695/Product: follitropin receptor #status predicted <SIG>
 F:16-366/Domain: extracellular hormone binding #status predicted <MAT>
 F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat #status predicted <EHB>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:367-387/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:398-421/Domain: transmembrane #status predicted <TM1>
 F:444-465/Domain: transmembrane #status predicted <TM2>
 F:486-508/Domain: transmembrane #status predicted <TM3>
 F:529-550/Domain: transmembrane #status predicted <TM4>
 F:574-597/Domain: transmembrane #status predicted <TM5>
 F:609-630/Domain: transmembrane #status predicted <TM6>
 F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 Query Match 12.5% Score 414.5, DB 1; Length 695;
 Best Local Similarity 25.0%; Pred. No. 2,2e-24;
 Matches 131; Conservative 94; Mismatches 233; Indels 67; Gaps 13;
 55 LKXDFQSLS-SLOALDSWNAIRSIHPEAF-----STLNSL----- 90
 Db IERNSEVGLSPFESVILMLNKNGIOEIHNCAGFNGTQDLDELNSDNNNEELPNDVFGASG 219
 QY 91 -VKLDLTNOQLTTLPLAGIGLMHKLKGNALSOAFSKDSFPKRLIEVYAAQCCPYG 149
 Db 220 PVILDISIRTHSLPSVLENLKLKRASTYVNLKPLLEKLVALMEASLTYPHCCAF- 278
 QY 150 MCASFRRASGWEADLH-----LDDESSRRPGLLARQENHYDDDLDELQ 198
 Db 279 -----AMNRROISELHPICNKSILRQEDVYMTQGTGSSLAEDNESSYSRGFMTYT 331
 QY 199 EMEDESKPFP--SVCCSPTPGPFKCEYLFESWGRILAVMAIVLSVLCNGLVLTTFVAG 256
 Db 332 EFYDLCNEVVDYTCSPKPAFNCEEDILGYNILRVLIWFSILATIGN-ITLVLTTS 390
 QY 257 PAPIPVKPYVGAAGANTLTGISCGLASVDALTGQSPSEYARWETGLCGRANGFLAV 316
 Db 391 QYKLYPFRFLMCMLAFADLCIGIYLLIASVDLHTKSQHNVAIDMGTAGDAGGFFTV 450
 QY 317 LGSEASVLLTLTAAVO-----CSVSVCAVRAKSPSLGVSAGVAGLGLAAGLAAALP 370
 Db 451 FASELSVYTLTATLERMTITTHAQDLDC-----KVQLRHAASVVMGWI-FAFAALFP 504
 QY 371 LASVGEYGAAPLCLPYAPPEGOPALGFTVALVMNSFCFLVNVAGATIKLYCDLPRGDF- 429
 Db 505 IEGISSTYKVSICLPM--DISPLISQLYVMSLVNLVAFVVICGYHITLYVRNPNIV 562
 QY 430 EAVWDCAVRHYAVMLIFADGLLYCPVAFLFASMLGEPVTPPEAVKSVLLVLDLPACLN 489
 Db 563 SSSSDPTRIAKHMAHMLITPDLICMAPISFPALISASLKVPLITVSKAKITLVLFPHINSCAN 622

QY 490 PLTYLLENFPRDDLRRLPRAGDSGLPAAVAAAGLEKSSCDST 533
 Db 623 PFLYALFTTNFRDRPFTLILSKG-----CYEQAOQIVRTERTST 661
 RESULT 10
 JN0898
 Follitropin receptor precursor - crab-eating macaque
 N:Alternate names: follicle-stimulating hormone receptor (FSHR)
 C:Species: Macaca fascicularis (Crab-eating macaque)
 C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text_change 10-Sep-1999
 A:Accession: JN0898, S36452
 R:Gromoll, J.; Dankbar, B.; Sharma, R.S.; Meschlag, B.
 A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of
 A:Reference number: JN0898; MIM:94071854; PMID:7504463
 A:Accession: JN0898
 A:Molecule type: mRNA
 A:Residues: 1-695 <GRO>
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CAA52463.1; PID:9396802
 A:Note: the authors translated the codon AGT for residue 488 as Arg
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follitropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotei
 F:1-171/Domain: signal sequence #status predicted <SIG>
 F:18-695/Product: follitropin receptor #status predicted <SIG>
 F:399-421/Domain: transmembrane #status predicted <TM1>
 F:444-465/Domain: transmembrane #status predicted <TM2>
 F:486-508/Domain: transmembrane #status predicted <TM3>
 F:529-550/Domain: transmembrane #status predicted <TM4>
 F:574-597/Domain: transmembrane #status predicted <TM5>
 F:609-630/Domain: transmembrane #status predicted <TM6>
 F:191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 Query Match 12.5% Score 414.5, DB 1; Length 695;
 Best Local Similarity 25.3%; Pred. No. 2,2e-24;
 Matches 134; Conservative 89; Mismatches 228; Indels 79; Gaps 13;
 55 LKXDFQSLS-SLOALDSWNAIRSIHPEAF-----STLNSL----- 90
 Db IERNSEVGLSPFESVILMLNKNGIOEIHNCAGFNGTQDLDELNSDNNNEELPNDVFGASG 219
 QY 91 -VKLDLTNOQLTTLPLAGIGLMHKLKGNALSOAFSKDSFPKRLIEVYAAQCCPYG 149
 Db 220 PVILDISIRTHSLPSVLENLKLKRASTYVNLKPLLEKLVALMEASLTYPHCCAF- 278
 QY 150 MCASFRRASGWEADLH-----LDDESSRRPGLLARQENHYDDDLDELQ 198
 Db 279 -----AMNRROISELHPICNKSILRQEDVYMTQGTGSSLAEDNESSYSRGFMTYT 331
 QY 199 EMEDESKPFP--SVCCSPTPGPFKCEYLFESWGRILAVMAIVLSVLCNGLVLTTFVAG 256
 Db 332 EFYDLCNEVVDYTCSPKPAFNCEEDILGYNILRVLIWFSILATIGN-ITLVLTTS 390
 QY 257 PAPIPVKPYVGAAGANTLTGISCGLASVDALTGQSPSEYARWETGLCGRANGFLAV 316
 Db 391 QYKLYPFRFLMCMLAFADLCIGIYLLIASVDLHTKSQHNVAIDMGTAGDAGGFFTV 450
 QY 317 LGSEASVLLTLTAAVO-----CSVSVCAVRAKSPSLGVSAGVAGLGLAAGLAAALP 370
 Db 451 FASELSVYTLTATLERMTITTHAQDLDC-----KVQLRHAASVVMGWI-FAFAALFP 504
 QY 371 LASVGEYGAAPLCLPYAPPEGOPALGFTVALVMNSFCFLVNVAGATIKLYCDLPRGDF- 429
 Db 505 IEGISSTYKVSICLPM--DISPLISQLYVMSLVNLVAFVVICGYHITLYVRNPNIV 562
 QY 430 EAVWDCAVRHYAVMLIFADGLLYCPVAFLFASMLGEPVTPPEAVKSVLLVLDLPACLN 489
 Db 451 FASELSVYTLTATLERMTITTHAQDLCKVHR-----HAASVVMGWI-FAF 498
 QY 365 LAAALPLASVGEYGAAPLCLPYAPPEGOPALGFTVALVMNSFCFLVNVAGATIKLYCDL 424

Db 499 AALPFLFGISSYMKVYSICLPM--DIDSPLSQLYVMSLVLNLAIVLCGGTTHLYLV 566
 QY 425 PRDGF-EAWWDCAMVRHVAWMLIFADGLLYCPVAFELSPASMLGFPTPAVKSVLLVLP 483
 Db 557 RNNVIVSSSDTITAKRMALITFDLCNAPISFFPAISAKLPITLVSKAKILLVLFVP 616
 QY 484 LPACINLTLFLFNPFRDRLRRLPRAGDSGPLVAAAGELKSSCDST 533
 Db 617 INSCANPFLAIFTKNRDFILLSKFG-----CYEMQAIYRTETSSP 661

RESULT 11

thyrotropin receptor precursor - dog
 A:Accession: A40077
 N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor
 C:Species: Canis lupus familiaris (dog)
 C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
 C:Accession: A40077; S06933
 R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; Van Sande
 Science 246, 1620-1622, 1989
 A>Title: Molecular cloning of the thyrotropin receptor.
 A:Reference number: A40077; MUID:90084524; PMID:2556796
 A:Accession: A40077
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PAA>
 A:Cross-references: GB:M29957; NID:g164098; PIDN:AAA30901.1; PID:g164099
 R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; van Sande
 Nucleic Acids Res. 17, 10493, 1989
 A>Title: Nucleotide sequence of the dog thyrotropin receptor cDNA.
 A:Reference number: S06933; MUID:90098886; PMID:2602159
 A:Accession: S06933
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PAA>
 A:Cross-references: EMBL:X17146; NID:9849; PIDN:CAA5026.1; PID:9850
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <MAT>
 F:21-764/Product: thyrotropin receptor #status predicted <MAT>
 F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
 F:201-225/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

Query Match 12.5%; Score 414.5; DB 2; Length 764;
 Best Local Similarity 23.7%; Pred. No. 2.5e-24;
 Matches 141; Conservative 102; Mismatches 258; Indels 95; Gaps 12;

QY 33 SLPAHPASIALAASNTTASGKLEDTFSQ-----SSLOALDLSWNA-IRSHPEAF 84
 Db 166 STPAN-----AFQGLCNFTLTLYKYNNGFTSIQGHAFNGTLDVAVLNKNKYLSAIDKDAF 221
 QY 85 STLHS-LVKLDLTDNQLTTLPLAGLGLIMHLKLGNAISQAFSKDSEPKLILEVPYAY 143
 Db 222 GGYVSGPTLDDISYTVLALPSKGLHLKELIARNMTWLTKRLPLSLFHLTRADLSYPS 281
 QY 144 QCCPY-----GMCASF-----FKASG 159
 Db 282 HCCAFKNOKKIRGITLESIMCNESIRSLRORRSVNTLNGPPDEYEELGDSHAGYDNS 341
 QY 160 QW-----EAEDLHDDSSSKRPLGLLARAQENHYDODLDELQLEMEDSKP 205
 Db 342 QPQDTSNSHYVFEFEDEDEITLFGQELKNPQETLOAFDNHSDYIV-----CGG 392
 QY 206 HPSVQCSPTPGPKPCPEYLEFSMGIRLAWAIVLISVLCNGLVLLTVFAGGAPADLPVKE 265
 Db 393 NEDWCTPKSDEFPNCEIDIMGYKFLRIYVWFSLLALDGNVFLVILVLSHYKLTVP-RF 451

QY 266 VVGAIGANTLTIGISGLLASVYALTFGQSEYGARWETIGCATGFLAVLGSEASVL 325
 Db 452 LMCNLAIFDFCGWVLLILLIASVLYTHSEYYNHAIDMIDTGGCMTAGFTVFASELSTYT 511
 QY 326 LTLAAQCSVSVSCYRANGKSPSLGVSAGVLCGLALGLAALPLASVGEYASPLCLP 385
 Db 512 LTVITLERVYATTFMRDRKIRLHAYAIWGWCVCCFLALLPLVQISSYAVSICLP 571
 QY 386 YAPBEGQALFETVALVMNNSFCFLVAGAYIKLYCDLPRGDE-AVWDCAMVRHVAWL 444
 Db 572 M--DIETPLATLITVLLNVAFLIYVSCVYKITYIRNPQWPGDKDKIAKRMAYL 629
 QY 445 IFADGLLYCPVAFELSPASMLGFPTPAVKSVLLVLPRLPCLPILLYLFPNFRDRL 504
 Db 630 IFTDFMCMAPISEFVALSALMNKPLITVNSKILLVLEFPLNSCANPFLAITTKAFQDY 689
 QY 505 RLRRPAGDSGPLVAAAG-----ELEKSSCDSTQALVAFSDVLLLENS 549
 Db 690 FILLSKFGICCKROAQRGQVRSFKNSAGIQIQKTRDKRQSLPMNODYEELLENS 745

RESULT 12

JC5643
 thyroid stimulating hormone receptor precursor - sheep
 C:Species: Ovis orientalis aries; Ovis ammon aries (domestic sheep)
 C>Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 20-Jun-2000
 C:Accession: JC5643
 R:Bockmann, J.; Winter, C.; Wilkowskl, W.; Kreutz, M.R.; Boeckers, T.M.
 Biochem. Biophys. Res. Commun. 238, 173-178, 1997
 A>Title: Cloning and expression of a brain-derived TSH receptor.
 A:Reference number: JC5643; MUID:97445147; PMID:9299474
 A:Accession: JC5643
 A:Molecule type: mRNA
 A:Residues: 1-764 <BOC>
 A:Cross-references: GB:Y13434; NID:92462632; PIDN:CAA73846.1; PID:92462633
 A:Experimental source: hypothalamus
 A:Note: the sequences of residues 4-7 and 8-17, 4136-439 and 440-449 are interchanged
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-764/Product: thyroid stimulating hormone receptor #status predicted <MAT>
 F:179-200/Domain: extracellular #status predicted <EXC>
 F:303-382/Domain: transmembrane #status predicted <TM1>
 F:416-439/Domain: transmembrane #status predicted <TM2>
 F:441-474/Domain: transmembrane #status predicted <TM3>
 F:495-518/Domain: transmembrane #status predicted <TM4>
 F:538-561/Domain: transmembrane #status predicted <TM5>
 F:581-604/Domain: transmembrane #status predicted <TM6>
 F:621-650/Domain: transmembrane #status predicted <TM7>
 F:661-684/Domain: transmembrane #status predicted <TM7>

Query Match 12.5%; Score 414.5; DB 2; Length 764;
 Best Local Similarity 25.0%; Pred. No. 2.5e-24;
 Matches 143; Conservative 93; Mismatches 244; Indels 93; Gaps 15;

QY 33 SLPAHPASIALAASNTTASGKLEDTFSQ-----SSLOALDLSWNA-IRSHPEAF 84
 Db 166 SVAN-----AFQGLSNFTLTLYKYNNGFTSIQGHAFNGKLDVAVLNKNKYITVDOAF 221
 QY 85 STLHS-LVKLDLTDNQLTTLPLAGLGLIMHLKLGNAISQAFSKDSEPKLILEVPYAY 143
 Db 222 AGYVSGPTLDDISYTVLALPSKGLHLKELIARNMTWLTKRLPLSLFHLTRADLSYPS 281
 QY 144 QCCPY-----GMCASF-KASGQW----- 161
 Db 282 HCCAFKNOKKIRGITLQSLMCNNSIWGLRORRSASALNPGYOEYEDLIGDSAGYKENS 341
 QY 162 EADDLH-----LDDSS-----KRLPGLLARAQENHYDODLDELQLEMEDSKP 205
 Db 342 KFDQTHNSHYVFEFEDEDEITLFGQELKNPQETLOAFDNHSDYIVGSGSEM----- 396
 QY 206 HPSVQCSPTPGPKPCPEYLEFSMGIRLAWAIVLISVLCNGLVLLTVFAGGAPADLPVKE 265

Db 397 ----VCTPKSDFNCDIEDIMYKFLRIYVWFVSLALLGNFVLVITLTSHYKLTVP-RF 451

QY 266 VCAIAGANTLTGISCGLASVDALITGQSEYCARMTGICGRATGFLAVLGSEAVYL 325

Db 452 LMCALRADPCMGILYLLINSVDLYTQSEYNNHADMOTGCGCTAGCFEYFASSELVYT 511

QY 326 LTLAAGCCSVSVSCVRAVYKGSFSLGSVAGVIGLALAGLAALPLASVGYGASPTCLP 385

Db 512 LTVITLERWYATFPAHMLDRKIRLMHAYVIMLGWCCFLALLPLVIGISSYAVSICLP 571

QY 386 YAPREGPALGFTVALVMNNSFCFLVYAGAYIKLYCDLPR----GDEFAVWDCAMVRH 440

Db 572 M--DTEPLALATITLLVLLNITIAFTIYACAYKITYTVRPNHYPGD---KDTRIAKR 625

QY 441 VAMLIFADGLLYCPVAFSLFASMLGLFPVTPPEAVKSVLVLVPLPACTNPLVLLFPHF 500

Db 626 MAVLITFDGCMAPISFYALSLAMNKRPLIVTNSKILLVLFYPLNSCANPFLYATFPAK 685

QY 501 RDLRLRPRAGDSGPLAYAAAGLEKSSCDST 533

Db 686 QRDVFMILSKFGICKRQAOAYRGO-RVSSKNST 717

RESULT 13

A42395

Lutropin receptor - mouse

N:Alternate names: luteinizing hormone-choriogonadotropin receptor

C:Species: Mus musculus (house mouse)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999

C:Accession: A42395

R:Guderamn, T.; Birnbaumer, M.; Birnbaumer, L.

J. Biol. Chem. 267, 4479-4488, 1992

A:Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenyly

A:Reference number: A42395; MUID:92165799; PMID:1311310

A:Accession: A42395

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-700 <GND>

A:Cross-references: GB:M81310; GB:M81318; NID:g198811; PIDN:AAA39432.1; PID:g198812

A:Note: sequence extracted from NCBI Backbone (NCBI:84064, NCBI:84066)

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C:Keywords: G protein-coupled receptor; transmembrane protein

F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 12.5%; Score 412.5; DB 2; Length 700;

Best Local Similarity 23.5%; Pred. No. 3.3e-24;

Matches 138; Conservative 108; Mismatches 267; Indels 73; Gaps 14;

QY 14 YRSGIPGSHASVRSOGLSLPAHPASLALASNTTASGKLENDPFSQLS-----L 66

Db 164 YTTIPENAP-----OGM-----NNESITLKLNGEVEVSHAFNGTTL 203

QY 67 QALDLSMNA-IRSHIPAFSTLSLVKLDLNDQTLTLLPLAGLGLHMLKGLALASQA 125

Db 204 ISLEKENVILEKMHSTFGATGATPSTLIDVSTKQLAPSHGLESTQTLTATSSYSIKTL 263

QY 126 FSDSPFKLILEVRYAYOCCPYGMCASFPGKASGWEAEHLHDDSESKRPILGLARQA 185

Db 264 PSREKFTSLVAVLTYPHSCAFRNLPK-----KEQNFSPISFENESKQCESTYREA 315

QY 186 ENH--YDQDDELQ-----EMEDSKPHBSVOCSPTRPGPKCEVYFESGIRLAWAIY 238

Db 316 NNPTLYSAIFENELSGMDYDYDFCSP-KTLQCTPPDPAFNPCEDIMGAYFLVLIWLIN 374

QY 239 LLSVLNGLVLLTVFAAGPAPLPVYFVVGAIAGANTLTGISCGLASVDALTFQGFSEK 298

Db 375 ILAIFGN-LTVLEVLLTSRYKLTVPREFLMCNLSFADPCMGILYLLIASVDSQGTQGYNNH 433

QY 299 GARMETGICGRATGFLAVLGSEAVYLTLTAOCCSVSVCAVYKGSFSLGSVAGVIG 358

Db 434 AIDMOTGCGCAAGFTYFASSELVYTITVITLERWYATFPAHMLDRKIRLMHAYVIMLG 493

QY 359 CLALAGLAALPLASVGEYASPLCLPYPAPREGPALGFTVALVMNNSFCFLVYAGAY 418

Db 494 GMIFSTIMATPLVGVSSYMKVSIQLP--DVESTLSQVYITLSTLLNAAVFAVYACAY 551

QY 419 KLYCDLPRGDEAV-WDCAMVRHAYMLIFADGLLYCPVAFSLFASMLGLFPVTPPEAVKSV 477

Db 552 KITFAVONPELTAPNKKDTKIAKMAILLITFDTCMADISFPAISAARKVPLITVNSKVL 611

QY 478 LTVLPLPACINPLVLLFNFHFRDRLRPRAGDSGPLAYAAAGLEKSSCDSTALV 537

Db 612 LVLEFPVNSCANPFLYATFPAKRDQDFLLLSRFG-----CCKHRAELY 655

QY 538 AFSVDLLILASERAGRPGLTYGFPs-----VTLASCOQPGAPRL 578

Db 656 RRKEFSACTPNSKNKGFPRSSK-----PSOAMKILSYHCQCPPTPPRV 697

RESULT 14

I48882

thyrotropin receptor precursor - mouse

N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor

C:Species: Mus musculus (house mouse)

C>Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 13-Aug-1999

C:Accession: I48882

R:Stein, S.A.; Oates, E.L.; Hall, C.R.; Grumbles, R.M.; Fernandez, L.M.; Taylor, N.A.

Mol. Endocrinol. 8, 129-138, 1994

A:Title: Identification of a point mutation in the thyrotropin receptor of the hyl/hy

A:Reference number: A44271; MUID:94224232; PMID:8170469

A:Accession: I48882

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-764 <RES>

A:Cross-references: EMBL:U02602; NID:9575923; PIDN:AAB60455.1; PID:9575924

C:Genetics:

A:Gene: TSHR

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea

C:Keywords: G protein-coupled receptor; transmembrane protein

F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:201-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 12.3%; Score 408; DB 2; Length 764;

Best Local Similarity 23.2%; Pred. No. 8.2e-24;

Matches 143; Conservative 99; Mismatches 251; Indels 124; Gaps 13;

QY 48 NTYASGKLENTSBSQ-----SSLOALDLSMNA-IRSHIPAFSTLSH-LVTKLDTDN 98

Db 177 NETLTLKLYNNGFTSVGHAFNGTKTDAVYLNNKKYLTALIDNDAFGVYSPTLIDVST 236

QY 99 QLTPLPAGLGLHMLKGLALASQAQFSPKSLILEVRYAYOCCPY-----GMC 151

Db 237 SYTALPESKGLHKLKELTAKDTWTLKLPDLSFLHLTRADSYPSHCACAFKNQKRGITL 296

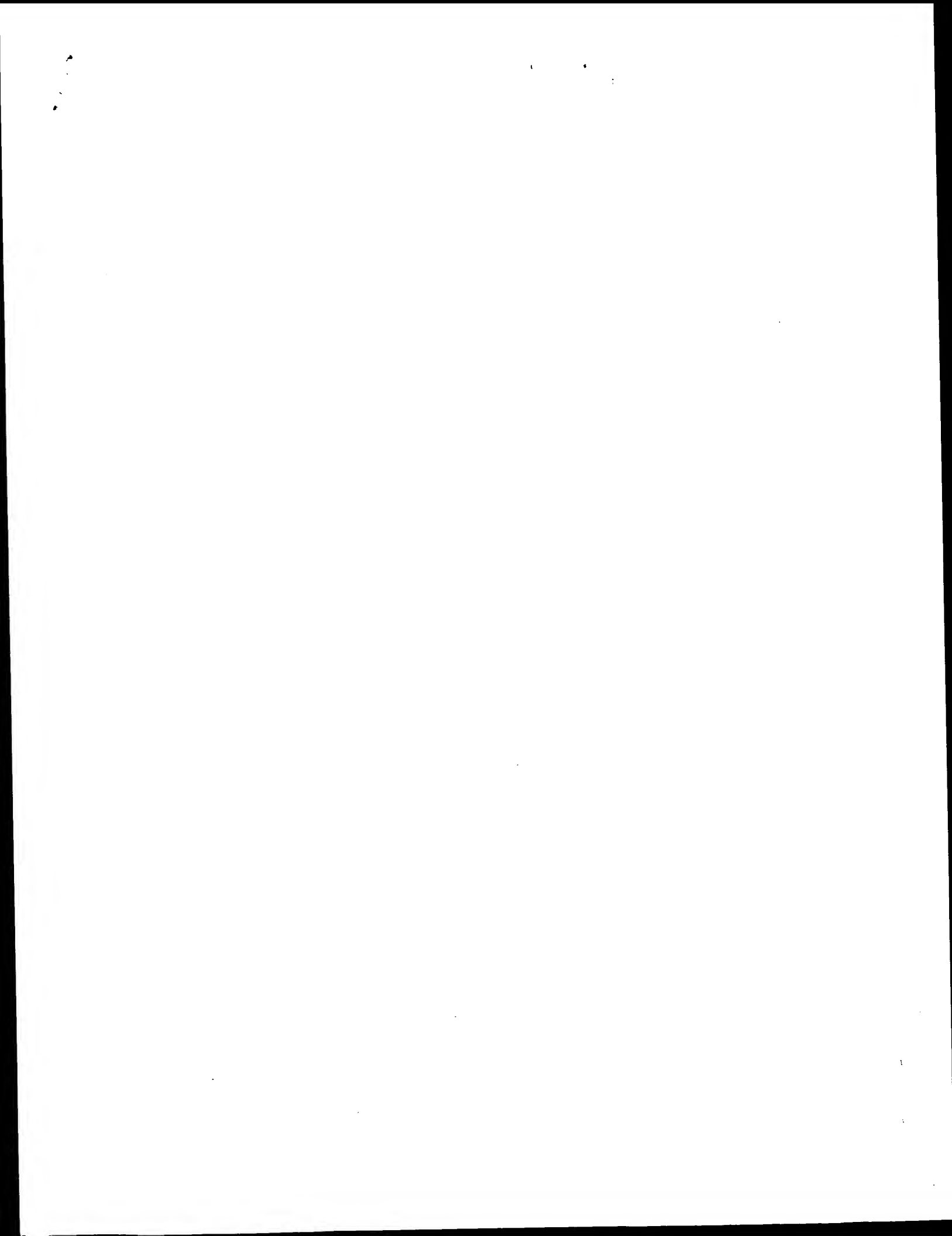
QY 152 ASF-----FKASGQ----- 161

Db 297 ESLMNCSSIRNLORKSVNLLNGPIYQVEEDPDGNSVGYKQNSKQESPSNHYVVF 356

QY 162 -EAEDLHDDSESKRPILGLARQAQENHVDQIDELQLEMEDSKPHBSVOCSPTRPGPKR 220

Db 357 EEOEDEYVGGQELKNPQEBTTLQAFESHYDY-----CGDNEDMVCPTPSDFNF 407

QY 221 CEYLFESWIGIRLAWAIVLLSVLCNGLVLLTVFAAGPAPLPVYFVVGAIAGANTLTGIS 280



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:34:40 ; Search time 7.3981 Seconds
(without alignments)
1233.683 Million cell updates/sec

Title: US-09-851-595-5
Perfect score: 3307
Sequence: 1 NTHYRESWACRYRSGIPG.....GGLSGGGGFGPSGLAFASHV 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 92612 seqs, 14418503 residues

Total number of hits satisfying chosen parameters: 92612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCRTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	29.0	631	10	US-09-862-767A-2
2	790	23.9	436	10	US-09-862-767A-4
3	783	23.7	431	10	US-09-862-767A-8
4	727.5	22.0	358	10	US-09-862-767A-6
5	539	16.3	861	10	US-09-804-551B-20
6	417	12.6	675	10	US-09-877-804-7
7	417	12.6	692	10	US-09-877-804-6
8	414.5	12.5	695	10	US-09-804-626-8
9	407.5	12.3	674	10	US-09-877-804-3
10	407.5	12.3	700	10	US-09-877-804-2
11	381	11.5	699	10	US-09-804-626-6
12	370.5	11.2	260	10	US-09-925-299-975
13	370.5	11.2	458	10	US-09-862-767A-9
14	350.5	10.6	359	10	US-09-862-767A-7
15	319	9.6	458	10	US-09-862-767A-5
16	316.5	9.6	694	10	US-09-928-175-8
17	316.5	9.6	730	10	US-09-928-175-7
18	309	9.3	718	10	US-09-928-175-3
19	309	9.3	754	10	US-09-928-175-2

20	306.5	9.3	718	10	US-09-928-175-21	Sequence 21, Appl
21	306.5	9.3	737	10	US-09-928-175-20	Sequence 20, Appl
22	302.5	9.1	646	10	US-09-928-175-13	Sequence 13, Appl
23	302.5	9.1	682	10	US-09-928-175-12	Sequence 12, Appl
24	269.5	8.1	757	10	US-09-928-175-24	Sequence 24, Appl
25	210.5	6.4	396	10	US-09-895-686-4	Sequence 4, Appl
26	159	4.8	270	10	US-09-764-855-145	Sequence 145, Appl
27	153.5	4.6	616	12	US-10-052-586-158	Sequence 158, App
28	150.5	4.6	164	10	US-09-764-853-1775	Sequence 775, App
29	149.5	4.5	80	10	US-09-925-301-1476	Sequence 1476, Ap
30	149	4.5	116	10	US-09-764-877-1993	Sequence 1993, Ap
31	145	4.4	78	10	US-09-925-299-10991	Sequence 1091, Ap
32	145	4.4	196	10	US-09-764-826-55	Sequence 55, Appl
33	142	4.3	26	10	US-09-864-761-42536	Sequence 42536, A
34	141.5	4.3	663	10	US-09-991-630-25	Sequence 25, Appl
35	141.5	4.3	2597	10	US-09-905-129-2	Sequence 2, Appl
36	141.5	4.3	2597	10	US-09-905-129-10	Sequence 10, Appl
37	141.5	4.3	2597	10	US-09-905-129-13	Sequence 13, Appl
38	141.5	4.3	2597	10	US-09-991-630-2	Sequence 2, Appl
39	141.5	4.3	2597	10	US-09-991-630-10	Sequence 10, Appl
40	141.5	4.3	2597	10	US-09-991-630-13	Sequence 13, Appl
41	140	4.2	673	9	US-09-992-598-52	Sequence 52, Appl
42	140	4.2	673	10	US-09-989-722-52	Sequence 52, Appl
43	140	4.2	673	10	US-09-782-980-59	Sequence 59, Appl
44	140	4.2	673	10	US-09-989-723-52	Sequence 52, Appl
45	140	4.2	673	10	US-09-989-279-52	Sequence 52, Appl

ALIGNMENTS

```
RESULT 1
US-09-862-767A-2
; Sequence 2, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MB10197-008P1RCP1CN(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127,856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054,646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-767A-2

Query Match          29.0%; Score 959; DB 10; Length 631;
Best Local Similarity 38.9%; Pred. No. 1.4e-67;
Matches 212; Conservative 88; Mismatches 189; Indels 56; Gaps 10;

QY      54 KLEKDTQSLSLQADLSMNAISHPKASHTSLKYLKDTLTDNLTPLAGLGIMH 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      59 QIKKGTGGLISLIDLSRMLIHHSRAFATLGTNTLDVSNELTSFTEGLINQ 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      114 LKLGKGNLALSGAFKDSFPKRILEVYAVOCPCYGCASFKASGQWEADLHDEES 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      119 LKLVGNKRLKALAKDPVNLRSLSVPYVCCAFWGCDSY----ANLNTEDNSLQDS- 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      174 SKRPLGLIARQENHYDDLDLQLEMEDSKRPS--VQSPTRGPCKPCYLFESWGIR 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      174 -----VAQKGTADANAVSTLENEE-HSQIITHCPSTGCAFKPCYLLGSWMIR 222
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      232 LAWAIVLVSVCGLVILTVFAGGAPALPVPKYVGAIAQANTLTGSCGLASVALT 291
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      223 LTVVFILVALFENLVLITLTFACSTL-TPSKLPIGLISVSNLFWGILYTGILTFDAVS 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

QY 292 FQOFEEYGRMETGCGCRATGFLAVIGSESVYLLTLAAVQCSVSVSCVAAKCSPTLSG 351
   : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Db 282 MGRFEPFGIMWETGSQCKKAGFLAVFSSSATFTLLMLATAYENLSKADIKNGKSHNLKO 341
QY 352 VRAGVUGCLIALAGLAALPLASVGEYGASPLCLPYAPPEGQPALGFTVALVMNNSCF 411
   : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Db 342 FRVAALLFLFGLATVAGCFFLFHNGEYSASPLCPF--PTGEMPSLTFTVLVLINSLAE 399
QY 412 VVAGAVYIKLYCPLPGDFEFVAWDCAMVRHVAWLIETFDGLLYCPAETSFASMLGLPYPT 471
   : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Db 400 LMAVYITKLYCKLEREDLSENSQSSMIKHWAMLIFNCLFPCVAFVFFSPALITAIISIP 459
QY 472 EAVKSVLVAVLPLPACLNPLLYLIRNPHRDD---LRRLRPARGSGPLAYAAAELEK 527
   | : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Db 460 EIMKSVTLTFEPLPACLNVLIVYFENPKREKMWLLKRRYTKRSGSEVSYSISQGGCLEQ 519
QY 528 -----SSCDSTQALVAFSDVD--LILEASEAGRPGLIETVYGFPSVTLI 568
   : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Db 520 DEYYDCGMVSHLQGNLTVCDDCESFLLTKRPSCKHILI-----KSHSCPALAVA 567
QY 569 SCOOP 573
   | | | | |
Db 568 SCQRP 572

```

```

RESULT 2
US-09-862-767A-4
; Sequence 4, Application US/09862767A
; Patent No. US20020034786A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
; FILE REFERENCE: MB101997-008P1RCP1CN1(M)
; CURRENT APPLICATION NUMBER: US/09/862,767A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/127, 856
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 60/054, 646
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-862-767A-4

```

	Query Match	23.9%	Score 790;	DB 10;	Length 436;
	Best Local Similarity	37.5%;	Pred. No. 1,5e-54;		
	Matches 174; Conservative	73;	Mismatches 157;	Indels 60;	Gaps 9;

QY	137	LEVPYAQCCEPYCMCASCFEFGKASGMEADLHLD---	EESKRRLGLLARQENHYDOD	192
Db	2	LSPVAAYCCAFMCDCSY----	ANLNINEDNLSLOCHSVAQEGTADAAVNTSTLENEHSQ	57
QY	193	IDELTQLEMEDSKPAPSVOCSEPTPGPKPCXEVLFEFSWGRILAVMAIVILSLVCNGIYLLTV	252	
		: :	:	
Db	58	I-----	IIHCPISTGAFCRCEYLSSWMIRLVMIIFVALFPNLVIILT	103
QY	253	FAGGPADLPVKRYVVGAIAGANTLTGTISCGLLASVDALTFGGQFSEYARGMETGLGCRAATG	312	
Db	104	FACSTSLPSKKLFITGLISTVSNLFEMGIYTGITFLIDAVSWGRFAEFGIMWETGSCKVAG	162	
QY	313	FLAVLGSEASVLLTTTLAAVOCSVSVCVRANGKPSLGSVAGVUGLCIALGLAALPLA	372	
Db	163	FLAVFSESIAIFLLMLATVERSISAKIMKNKGNHLKOFVAAALLAFLGTAVAGCPFLF	222	
QY	373	SVEGYASPLCLPEYPAPPEGOPALGFTVALVYMNSFCFLVVAGAYIKLYCDLPGRDFEAV	432	
Db	223	HREGEYSPLCLPE--PTGETPSIGFTVTIVLNLSTLAFLMAVITYTKLYCLEKDEDSN	280	
QY	433	WDCAVHVHVMILFEAGLLCTCPAVALFSAAMLGLFPVTPAEVRSVLLVLEPLACLNFLL	492	
		: : : : : : : : : : : : :		

```

Db      281  SSSSIKIKHAMVILFTNCIFCPVPVAFSPAPLITAISSPEIMKSVTLIFPLPACLNPLY 340
Qy      493  YLLFNPVHRDD-----LRRLRAPDSDGLPAAVAAAGLEK-----SSCD5 532
Db      341  YVFNPVKKEEDMKLLKRRVTKTKSSSVSVSISSGGCCEQDFYIDCGMYSHLGNLNLYCDC 400
Qy      553  TQALVAEFDV-----LLEASAGRPGLGTYGFSVTLISCOOP 573
Db      401  CESFLTKPVSCKHIL-----KSHSCALAAVASCQRP 432

```

```

RESULT 3
US-09-862-767A-8
: Sequence 8, Application US/09862767A
: Patent No. US20020034786A1
: GENERAL INFORMATION:
: APPLICANT: Pan, Yang
: TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
: FILE REFERENCE: MB01897-00861RCP1CN1(M)
: CURRENT APPLICATION NUMBER: US/09/862,767A
: CURRENT FILING DATE: 2001-05-21
: PRIOR APPLICATION NUMBER: US 09/127,856
: PRIOR FILING DATE: 1998-08-03
: PRIOR APPLICATION NUMBER: US 60/054,646
: PRIOR FILING DATE: 1997-08-04
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8
: LENGTH: 431
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-862-767A-8

```

Query Match	23.7%	Score 783;	DB 10;	Length 431;
Best Local Similarity	37.4%;	Pred. No. 5.3e-54;		
Matches 173;	Conservative 73;	Mismatches 157;	Indels 60;	Gaps 9

QY	137	LEVPAYOCCPYGMCASFEMKASGQWEADLHLD---	EESKRPLGLDARQAEHHYDD	192
Db	2	LSVPAYOCCAFWGDYSY----	ANNTEDNSLQDHSVQAEKGTADANVSTYLENEHSQ	57
QY	193	LDLQLEMEDSKPHPSVOCSPPTGPPEKPCCEYLFE	SMGRFLAMVALYLLSVLONGLVLT	252
Db	58	I-----	IHCPTSTGAFKFCQETYLGSMMRLTYWFLFLVALFENLVLT	103
QY	253	FAGGAPRLPYKPVVGATAGANTLLGISCGLASVADL	TEFGQFSEYGARMEFTGLGCRATG	312
Db	104	FASCTSPSSKLEFLGLISVGNLFNGIYTGILTF	LDVAWSGFRFAFGIMWEGSCOKVAG	162
QY	313	FLAVLGSASVLLTLTAAVOCVSYSVCRA	TKSPKSSLSGVSRAVIGTCLAGLAALPLA	372
Db	163	FLAVSSSSALFELMLAVERSLSAKDI	TMKNKGSNHLQFRVALALAEFLGATVACCFPLF	222
QY	373	SVGEYGAASPLCLPAPPEGGQPAALGFTYAL	VMNNSFCFLVAVAGYIKFYCDLPRGDEAV	432
Db	223	HREGYSPCLCLPR--	PLGTERPSIGFTYVLLNSLATLMAVITYTKLCNIEKEDLSEN	280
QY	433	WDCAMRHVAMLLFADGLLYCPVAPLFS	ASMUGLFPVTPPEAVKSVLLVPLPLACINBL	492
Db	281	SOSSMIKHVAMLLFTNCFICFCPVAF	FFSFAPLITIAISIEPEIMKSVTLTFPLPACINBL	340
QY	493	YLLENPHFRD---	LRLRPAGDSGLAVAAAGLEL	532
Db	341	YVFPNPFKEDMKLLKRRVTKKSGSV	SVSISSQGGCEQDDPYDQGMVSHLQGNLTVCDC	400
QY	533	TQALVAFSDV--	LILEASEAGRPGLTETGFP	SVTLISCOQ 572
Db	401	CESFLLTKPVSKHLI-----	KSNRSCALAVASQR	431

RESULT 4

US-09-862-767A-6

Sequence 6, Application us/09862767A

Query Match

12.64; Score 417; DB 10; Length 675;

Best Local Similarity 25.0%; Pred. No. 4.7e-25;
Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

```

QY 16 SGIPSTHASYERSOGL-----SLPAHPASIALAASNTASGKLEXTDFSSLSL 66
Db 75 SNLKLHEIRLEKANNLLIYNPEAFQNLPSLRVLLISNT--GIRKLPAVHKIQSL 127
QY 67 Q-----ALDSMNAIRSIHPEAF-----S 85
Db 128 OKVLIDIDNINIHIVARNSEFMGLSFESVILWLSKNGIEIHNCANFGTQDELNLSDNN 187
QY 86 TLHSL-----VKLDLTDNOLITPLLAGLGLMLKLGKGNLALSOARSKDSFPL 134
Db 188 NLEELPNDVFOGASGPVILIDISTRTVHSLPNHGLENLKRLARSTYRLKLPNDKEVTL 247
QY 135 RILEVPAVYOCPPYGCASCFKASGQWEADLH-----LD 169
Db 248 MEASLTPYSHCCAF-----ANLKRQISLHPICNKSILRQDIDMTQIGDQVSLID 239
QY 170 DEESKRPLGLARQAEHNYDODLELOEMEDSKPHPSVOCSPTPGPKFCEYLFESWG 229
Db 300 DEPS---YKGSDDMYNEFDYLCN---EYVD-----VTCSPKPAFNPCEIDINGYNI 346
QY 230 IRLAWMAIVLISVLCNGLVLLTVFAGGAPLPVKEVVGAIAGANTLTGICGLASVDA 289
Db 347 LRVLMFTISLITAGNTVLV-VLTTSQYKLTVPRLMGNLAFADLCIGIYLLIASVDI 405
QY 290 LTFQGFSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVQ-----CSVSVSCVRAY 343
Db 406 HTSQYHNHAYLDMOTGAGCDAGFFVFASLSVTLTATLERMTTHHAMQEC-----461
QY 344 GKSPSLGSVRAGVLCGLALAGLAAALPLASVGEYASPLCLPYAPPEOPALGFTVALV 403
Db 462 -KVOLRHAASVWVLG-WTFAPAAALPFIETGSSYMKVSIQCPM--DIDPSLSQLYVALL 517
QY 404 MNNSCFELVAGAYIKLYCDLPRGDF-EAVWDCAMVRHVAALIFADGLICPAVLESPAS 462
Db 518 VLVNLAFAVYICGCTHYIYLVNPNPTIVSSSDTKIAKRMATLFTDLCAPISFFAISA 577
QY 463 MGLFPVPEAVKSVLLVLPPLACNLPLVLLFNPFRDLRLRRRAGDSGLAYVAA 522
Db 578 SLKVPPLITVSRAKILLVLFYPIVNSCANPFLYALFTKFRDFEIL-----622
QY 523 GELEKSSCDSTQALVAFSDVDLLEASAGRPGLFTYGPFSVTLISCOOPGAPRLGESH 582
Db 623 --LSKFGCYEMAOIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTVNSY 664
QY 583 CVEPEGNH 590
Db 665 VLVPL-NH 671

```

RESULT 7
US-09-877-804-6
Sequence 6, Application US/09877804
Patent No. US20020061557A1
GENERAL INFORMATION:
APPLICANT: McFarland, Karoly
APPLICANT: Nikolics, Karoly
APPLICANT: Segalo, Deborah L.
APPLICANT: Seeburg, Peter H.
TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
FILE REFERENCE: P0576P1C2
CURRENT APPLICATION NUMBER: US/09/877,804
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 08/207,814
PRIOR FILING DATE: 1994-03-07
PRIOR APPLICATION NUMBER: US 07/781,153
PRIOR FILING DATE: 1991-10-31
PRIOR APPLICATION NUMBER: US 07/347,683
NUMBER OF SEQ ID NOS: 22
SEQ ID NO 6

LENGTH: 692
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: deduced sequence
US-09-877-804-6

Query Match 12.6%; Score 417; DB 10; Length 692;
Best Local Similarity 25.0%; Pred. No. 4.8e-25;
Matches 167; Conservative 86; Mismatches 251; Indels 164; Gaps 23;

```

QY 16 SGIPSTHASYERSOGL-----SLPAHPASIALAASNTASGKLEXTDFSSLSL 66
Db 92 SNLKLHEIRLEKANNLLIYNPEAFQNLPSLRVLLISNT--GIRKLPAVHKIQSL 144
QY 67 Q-----ALDSMNAIRSIHPEAF-----S 85
Db 145 OKVLIDIDNINIHIVARNSEFMGLSFESVILWLSKNGIEIHNCANFGTQDELNLSDNN 204
QY 86 TLHSL-----VKLDLTDNOLITPLLAGLGLMLKLGKGNLALSOARSKDSFPL 134
Db 205 NLEELPNDVFOGASGPVILIDISTRTVHSLPNHGLENLKRLARSTYRLKLPNDKEVTL 264
QY 135 RILEVPAVYOCPPYGCASCFKASGQWEADLH-----LD 169
Db 265 MEASLTPYSHCCAF-----ANLKRQISLHPICNKSILRQDIDMTQIGDQVSLID 316
QY 170 DEESKRPLGLARQAEHNYDODLELOEMEDSKPHPSVOCSPTPGPKFCEYLFESWG 229
Db 317 DEPS---YKGSDDMYNEFDYLCN---EYVD-----VTCSPKPAFNPCEIDINGYNI 363
QY 230 IRLAWMAIVLISVLCNGLVLLTVFAGGAPLPVKEVVGAIAGANTLTGICGLASVDA 289
Db 364 LRVLMFTISLITAGNTVLV-VLTTSQYKLTVPRLMGNLAFADLCIGIYLLIASVDI 422
QY 290 LTFQGFSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVQ-----CSVSVSCVRAY 343
Db 423 HTSQYHNHAYLDMOTGAGCDAGFFVFASLSVTLTATLERMTTHHAMQEC-----478
QY 344 GKSPSLGSVRAGVLCGLALAGLAAALPLASVGEYASPLCLPYAPPEOPALGFTVALV 403
Db 479 -KVOLRHAASVWVLG-WTFAPAAALPFIETGSSYMKVSIQCPM--DIDPSLSQLYVALL 534
QY 404 MNNSCFELVAGAYIKLYCDLPRGDF-EAVWDCAMVRHVAALIFADGLICPAVLESPAS 462
Db 535 VLVNLAFAVYICGCTHYIYLVNPNPTIVSSSDTKIAKRMATLFTDLCAPISFFAISA 594
QY 463 MGLFPVPEAVKSVLLVLPPLACNLPLVLLFNPFRDLRLRRRAGDSGLAYVAA 522
Db 595 SLKVPPLITVSRAKILLVLFYPIVNSCANPFLYALFTKFRDFEIL-----639
QY 523 GELEKSSCDSTQALVAFSDVDLLEASAGRPGLFTYGPFSVTLISCOOPGAPRLGESH 582
Db 640 --LSKFGCYEMAOIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTVNSY 681
QY 583 CVEPEGNH 590
Db 682 VLVPL-NH 688

```

RESULT 8
US-09-804-626-8
Sequence 8, Application US/09804626
Patent No. US20020128190A1
GENERAL INFORMATION:
APPLICANT: Lobel, Leslie
APPLICANT: Joyce
TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN
FILE REFERENCE: 0575/62259/JPM/SHS
CURRENT APPLICATION NUMBER: US/09/804,626
NUMBER OF SEQ ID NOS: 8

```

; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-804-626-8

```

```

Query Match      12.5%; Score 414.5; DB 10; Length 695;
Best Local Similarity 25.1%; Pred. No. 7.6e-25;
Matches 123; Conservative 94; Mismatches 231; Indels 43; Gaps 11;

```

```

QY 64 SSIALDLSS-WNIRSHPEASTLSHIVKLDLTDNQLTLPAGLGGLMLKLGMLAL 122
DB 193 TQIDAVNLSDNNLEELPNDVFHAGSPVLLDLSRRIRHSLPVGLEMLKKLRARSTYNL 252
QY 123 SQAFKSDPFKRLILEVPYAVOCCPYGMCASFRAQSCQMAEDLH-----LDDE 171
DB 253 KRLPTLEKLVALMEASLTYPSSHCCAF-----ANWRQISELHPICNLSILROEVDYM 304
QY 172 ESKRRLGLARQAEHNYDDDELQLEMEDSKRHP--SVQSPPTGPPKPCXYLFESWG 229
DB 305 TOARGORSSLAEDNESSYSGFDYTYTEFDYDLCNEVVDYTCSPKPAFNCEIDMGYNI 364
QY 230 IRLAVMAIVLTVLCNGVLLTVFAGGPAPLPVKFYVGAIGANTLTGISCGLLASYDA 289
DB 365 LRVLIWIFISILATIGN-ILYVILITTSQYKLTVPRLMCLNAFADLCIGIYLLIASYDI 423
QY 290 LTFGESEYGARWETGICRATGFLAVLGSEASVLLTLAAV-----GCVSVSCVRAV 343
DB 424 HTKQYHNYALDMDGTAGCDAAFEFTVFASELVYTLATLEHMTTHIAMQDLC----- 479
QY 344 GKSSLSQSVRAGVGCALAGLAAALPLASVGEYGASPLCPYAPPEGAPALGFTVALV 403
DB 480 -KVQRLRAASAVWGMW-FAPRAALFPIFGISYMKVSIICLPM--DIDSPLSOLIVMSLL 535
QY 404 MMSNFCFLVAGAYTKLYCDLPRGDF-EAVWDCANVRHVAAILFADGLLXCPVALSLAS 462
DB 536 VLNVIAFVIGCVYHITLYFRNPNIYSSSDTRIAKRAMAILFDLCMAISFALISA 595
QY 463 MGLGFPVPEAVKSVLLVPLPACLNPLLYLFPNHRDRLRLPRASGSPPLAYAAA 522
DB 596 SLKVPALTVSAKILLVLFHPINSCANPFLVATFKNRRDRDFILLSKCG-----CYEMQ 650
QY 523 GELEKSCDST 533
DB 651 AQIYRTETST 661

```

```

RESULT 9
US-09-877-804-3
; Sequence 3, Application US/09877804
; Patent No. US20020061557A1
; GENERAL INFORMATION:
; APPLICANT: Nikolic, Karoly
; APPLICANT: McFarland, Keith C.
; APPLICANT: Segalo, Deborah L.
; APPLICANT: Seeburg, Peter H.
; TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
; FILE REFERENCE: P0576P1C2
; CURRENT APPLICATION NUMBER: US/09/877,804
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 08/207,814
; PRIOR FILING DATE: 1994-03-07
; PRIOR APPLICATION NUMBER: US 07/781,153
; PRIOR FILING DATE: 1991-10-31
; PRIOR APPLICATION NUMBER: US 07/347,683
; PRIOR FILING DATE: 1989-05-05
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 3
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:

```

```

; OTHER INFORMATION: deduced sequence
US-09-877-804-3

```

```

Query Match      12.3%; Score 407.5; DB 10; Length 674;
Best Local Similarity 24.0%; Pred. No. 2.6e-24;
Matches 135; Conservative 99; Mismatches 254; Indels 75; Gaps 13;

```

```

QY 47 SNTTASGKLEXPSPQLSS-----LQALDLSMNA-IRSHPEASTLSHIVKLDLTDN 98
DB 151 NNSSVTLTKYNGEVEESHAFNGTTLISLEKENITLKHSHGAFGAGPGLDLSST 210
QY 99 QLTTLPLAGLGGLMLHLKGLALSQAFKSDPFKRLILEVPYAVOCCPYGMCASFRAK 158
DB 211 KLOALPSHGLESITLALSYSLKTLPSKEKFTSLVATLTYSHCCAFRNLPK----- 265
QY 159 GQWEAEHLHDEESSKRPGLLARQAEH--YDQDLEQLEMED-----SKRHPYQC 211
DB 266 ---KEQNFPSFISFENFSKQCESTVRKADNETLYSAIFEENELSGMDVDYGRCSF-KTLQC 321
QY 212 SPPTGPPKPCXYLFESWGIRLAVMAIVLTVLCNGVLLTVFAGGPAPLPVKRYVGAIA 271
DB 322 APEPDAPNCCEDINGYAFRLVILMLTNILALFGN-LTVLFVLLISRYKLTVPRLMCLNS 380
QY 272 GANTLTGISCGLASVDAITFGQSEYGARWETGICRATGFLAVLGSEASVLLTLAAV 331
DB 381 FADFCMGILYLLIASVDSOTKGOYNNHAIMDGTSGCGAAGFTVFASELVYTLVITL 440
QY 332 QCSVSVSCVRAVYKSPSLGSRVAGVLCALAGLAAALPLASVGEYGASPLCLPYAPREG 391
DB 441 ERWHTITVAVQDQKRLRHAIPIMLGWMLSTLTATMTLVGISYMKVSIICLPM--DYE 498
QY 392 QPALGFTVALVMMNSFCFLVAGAYTKLYCDLPRGDFEAV-WDCANVRHVAAILFADGL 450
DB 499 STLQSVYLLSTILLVNAFVIGACACIRITFEVQNPBELTAPRKDKRIAKMALLFTDPT 558
QY 451 LYCPAFLSFASMLGLFPVTPAVKSVLLVPLPACLNPLLYLFPNHRDRLRLR 510
DB 559 CMAISFALISAAPVPLITVNSKILLVLFYVNSCANPFLVATFKRQDFILLISR 618
QY 511 AGDSGPLAVAAAGLEKSCDSTQALVAFSDVLLLEASEACRPGLEY-----GFP- 563
DB 619 FG-----CCKRRA-----ELYRRKEFSYTNCKNKGFPG 647
QY 564 -----SVTLISQCGQGAPR 577
DB 648 ASKPSQATLKLTSTVHQQPLPPR 670

```

```

RESULT 10
US-09-877-804-2
; Sequence 2, Application US/09877804
; Patent No. US20020061557A1
; GENERAL INFORMATION:
; APPLICANT: Nikolic, Karoly
; APPLICANT: McFarland, Keith C.
; APPLICANT: Segalo, Deborah L.
; APPLICANT: Seeburg, Peter H.
; TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
; FILE REFERENCE: P0576P1C2
; CURRENT APPLICATION NUMBER: US/09/877,804
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 08/207,814
; PRIOR FILING DATE: 1994-03-07
; PRIOR APPLICATION NUMBER: US 07/781,153
; PRIOR FILING DATE: 1991-10-31
; PRIOR APPLICATION NUMBER: US 07/347,683
; PRIOR FILING DATE: 1989-05-05
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 2
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:

```


OTHER INFORMATION: deduced sequence
US-09-877-804-2

Query Match 12.3%; Score 407.5; DB 10; Length 700;
Best Local Similarity 24.0%; Pred. No. 2.7e-24;
Matches 133; Conservative 99; Mismatches 254; Indels 75; Gaps 13;

```

QY 47 SNTTASGRLEKDFEFSQS-----LQALDLSWNA-IRSHPEAFSTLHSLVKLIDLTGN 98
DB 177 NNESTLKLKXNGFEEVQSHAFNGTTLISLEKNEYIEKMHNGAFQCATGFSILDIST 236
QY 99 QLTPLPLAGLGLMHLKGLKGLASQAFSKDSFPRKILEVPYAYQCCPYGCASFRRAS 158
DB 237 KLAALPSHGLSISQTLALSSYSLSKTLPSKEKFTSLVATLTPSHCCAFRLPK----- 291
QY 159 GOWEADLHLDDESSKRPLGLLARQENH--YQDDDELQEMD-----SKPSPVQC 211
DB 292 ---KEQNFSTSIENFNSKQCESTYRKADNETLYSAIFEENELSGWDYDYGFCSP-KTLOQC 347
QY 212 SPTPPGPKPEEYLFESGIRLAWAIVLLSVLCNGLVLLTFVAGGPAPLPVKFVGAIA 271
DB 348 APBPDAFNPCEIDIMGYAFELVILINILAIFGN-LIVLFVLLTSRYKLTVPRLMCNLS 406
QY 272 GANTLIGISGLASVDALFFGOFSEYGARWETGLGRATGFLAVLGSESVLLITLAAV 331
DB 407 FADFCMGLVLLILASVDSQTKGQYINHAIDWQSGCGAGFTVFASFELSVYTLTVITL 466
QY 332 QCSVSVCVRANKSPSLGSRVAGVLCGLALAGLAALPLASVGEYGASPLCLPYAPPEG 391
DB 467 ERWHTTIVAVLDQKRLRHAIRLIMLGWLFSTLIATMPLVGISNMYKVSICLPM--DVE 524
QY 392 QPAALGFVALVMNSECFLVAVGAYIKLYCDLPRGDFEAV-WDCAMVRVAMLIADGL 450
DB 525 STLQVYILITILINVAFFIICACIKITFAVRNDELMTNDRKAKKMAILIFDFDT 584
QY 451 LYCPVAFLSFASMLGLPVTPEAVKSVLLVLPPLACPLNPLLYLFPNHRDRLRLRPR 510
DB 585 CMAPISFPAISAFKVPILITVNSKILLVLFYVNSCANPFLVAIFKAKQORDLLLSR 644
QY 511 AGDSGFLAVAAAGLEKSCDSQALVAVSDVLLLEASGAPPGLEYT-----GFP- 563
DB 645 FG-----CCKRRA-----ELYRRKEFSAVTSNCKNGFPG 673
QY 564 -----SVTLISCOQGPAPR 577
DB 674 ASKPSQATLKLTSTVHCQDPPIPR 696

```

RESULT 11

US-09-804-626-6
Sequence 6, Application US/09804626
Patent No. US20020128190A1
GENERAL INFORMATION:
APPLICANT: Lobel, Leslie
TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN OF
FILE REFERENCE: 0575/62259/JPW/SHS
CURRENT APPLICATION NUMBER: US/09/804,626
CURRENT FILING DATE: 2001-03-09
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin version 3.0
SEQ ID NO 6
LENGTH: 699
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-804-626-6

Query Match 11.5%; Score 381; DB 10; Length 699;
Best Local Similarity 25.3%; Pred. No. 3.2e-22;
Matches 122; Conservative 91; Mismatches 241; Indels 28; Gaps 10;
47 SNTTASGRLEKDFEFSQS-----LQALDLSWNA-IRSHPEAFSTLHSLVKLIDLTGN 98

```

DB 173 NNESTLKLKXNGFEEVQSHAFNGTTLISLEKNEYIEKMHNGAFRQATGPKTIDIST 232
QY 99 QLTPLPLAGLGLMHLKGLKGLASQAFSKDSFPRKILEVPYAYQCCPYGCASFRRAS 158
DB 233 KLAALPSHGLSISQTLALSSYSLSKTLPSKEKFTSLVATLTPSHCCAFRLPK----- 285
QY 159 GOWEADLHLDDESSKRPLGLLARQENH--YQDDDELQEMD-----SKPSPVQC 211
DB 286 PTKQNFSTSIENFNSKQ-CESTYRKANNKTLYSMALESLSGWDYDYGFCPL-KTPRC 343
QY 212 SPTPPGPKPEEYLFESGIRLAWAIVLLSVLCNGLVLLTFVAGGPAPLPVKFVGAIA 271
DB 344 APBPDAFNPCEIDIMGYAFELVILINILAIIGN-MTVLFVLLTSRYKLTVPRLMCNLS 402
QY 272 GANTLIGISGLASVDALFFGOFSEYGARWETGLGRATGFLAVLGSESVLLITLAAV 331
DB 403 FADFCMGLVLLILASVDSQTKGQYINHAIDWQSGCGAGFTVFASFELSVYTLTVITL 462
QY 332 QCSVSVCVRANKSPSLGSRVAGVLCGLALAGLAALPLASVGEYGASPLCLPYAPPEG 391
DB 463 ERWHTTIVAVLDQKRLRHAIRLIMLGWLFSTLIATMPLVGISNMYKVSICLPM--DVE 520
QY 392 QPAALGFVALVMNSECFLVAVGAYIKLYCDLPRGDFEAV-WDCAMVRVAMLIADGL 450
DB 521 STLQVYILITILINVAFFIICACIKITFAVRNDELMTNDRKAKKMAILIFDFDT 580
QY 451 LYCPVAFLSFASMLGLPVTPEAVKSVLLVLPPLACPLNPLLYLFPNHRDRLRLRPR 510
DB 581 CMAPISFPAISAFKVPILITVNSKILLVLFYVNSCANPFLVAIFKAKQORDLLLSR 640
QY 511 AG 512
DB 641 FG 642

```

RESULT 12

US-09-925-299-975
Sequence 975, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 975
LENGTH: 260
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (212)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-975

Query Match 11.2%; Score 370.5; DB 10; Length 260;
Best Local Similarity 36.3%; Pred. No. 6e-22;
Matches 78; Conservative 41; Mismatches 59; Indels 37; Gaps 5;
QY 382 LCLPYAPPEGQPAALGFVALVMNSECFLVAVGAYIKLYCDLPRGDFEAVWDCAMVRV 441
DB 1 LCLPYF--PGEFSPISLGFYLVILNLSLAFILMAVIYIKLYCNLEKEDLSNGSGSMIKHV 58
QY 442 AMLIFADGLYCPVAFLSFASMLGLPVTPEAVKSVLLVLPPLACPLNPLLYLFPNHR 501
DB 59 AMLIFNCLIFCPVAFVAFSPFLTAISISPEIKMSVYLIFPPLACPLNPLVLYVFENPKF 118

QY 502 DD-----LRLAPRAGDSGLAYAAAGELK-----SSCDSTQALVAFSD 541
 Db 119 EDMKLKRRVTKKSGSVSSISOGGLEQDFYDCGMYSHLQGNLTVCDCESFLTKP 178
 QY 542 VD---LLELSEAGRPGLTETGPPSVTLISCOOP 573
 Db 179 VSCKHLT-----KSHSCPALAVASCORP 201

RESULT 13

US-09-862-767A-9
 ; Sequence 9, Application US/09862767A
 ; Patent No. US20020034786A1

GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCPICN1(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/127,856
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 458
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-862-767A-9

Query Match 11.2%; Score 370.5; DB 10; Length 458;
 Best Local Similarity 25.5%; Pred. No. 1.2e-21;
 Matches 113; Conservative 80; Mismatches 200; Indels 51; Gaps 11;

QY 119 NLAISOAFKSDSPFKRLIE-----VPYAYCCPYGKASFFKSGQWEADLH-- 167
 Db 3 NLEKLARSTYNNLEKLPTEKLVALMEASLTYPHSCAF-----ANMKRQISELHP 54
 QY 168 -----IDDESSKRPGLLARQENHYDQDLDELQLEMEDSKPH--SVQCSPTPG 216
 Db 55 CNKSLTRQEVDTYQTRGQSSSLAEDNESSYRGFDWTFEFDYDLCEVVDVDCSPKPD 114
 QY 217 PERPCYELFESMGIRLAWMAIVLTVLCNGVILTVFAGGAPLPVYKFGVAGALAGANTL 276
 Db 115 AENPCDINGYNLRLVLPWITSLITLITGN--IIVLVITTSQYKLTVPFELMCNLAFADEL 173
 QY 277 TGISGGLASVDALTFQGESEYGARMETGLGCRATGFLAVLGSEASVLLTLTAAVQ--- 332
 Db 174 TGIYLLILASVDIHTKSYHNYAIDWQTGAGCDAAGFTVFASLSVYTLTATTLERWHT 233
 QY 333 --CSVSVGCVRAGKSPSLGSAVAGVLCALAGLAALPLASGEYGASPLCLPYAPPE 390
 Db 234 ITHAMQDLC-----KVQLRHAASVMWGMW-FAFAALFFIFGISYMKVSIQCPM--DI 285
 QY 391 GPPAALGFTVALVWMSFCLVAVAGAYIKLYCDLPBGDF--EAVWDCAMVRHVALITADG 449
 Db 286 DPLSOLLYMSLLVLVLAFFVIGCYIHLYLVARNNIVSSSDTIARMAALITPTDF 345
 QY 450 LLYCVAFILSFASMLGLFVTPPEAVSVLLVLPACNLPLLYLLENPHFRDRLRLRP 509
 Db 346 LCMAPISFPAISAKLVPLITVSKAKILLVLFHPINSCANPFLYAIPTKMRDRDFILLS 405
 QY 510 RAGDSGLAYAAAGELKSSCDST 533
 Db 406 KCG-----CYEMQAOIYRIETST 424

RESULT 14

US-09-862-767A-7
 ; Sequence 7, Application US/09862767A
 ; Patent No. US20020034786A1

GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCPICN1(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/127,856
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 359
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-862-767A-7

Query Match 10.6%; Score 350.5; DB 10; Length 359;
 Best Local Similarity 27.5%; Pred. No. 3.4e-20;
 Matches 99; Conservative 68; Mismatches 178; Indels 15; Gaps 6;

QY 175 KRLPLGLARQENHYDQDLDELQLEMEDSKPHPSVQCSPTPGPCEYLFESMGIRLAY 234
 Db 4 KNPQETTLQAFDSHYDVTI-----CGDSE--DMVCTPKSDEFNCPEDIMGYKFLRIV 54
 QY 235 WAIVLLSVLCNGVILTVFAGGAPLPVYKFGVAGALAGANTLTGISGGLASVDALTFQ 294
 Db 55 WFSYSLALGNVFLVLTLSHKLWVPRFLMCNLAFADECKMAYLLILASVDLYHSE 113
 QY 295 FSEYGARMETGLGCRATGFLAVLGSEASVLLTLTAAVQCSVSVAVRAGKSPISGVRA 354
 Db 114 YNNHAIDWQTGPGCNAGFTVFASLSVYTLTVTERWALIFAMRDLKILRHACA 173
 QY 355 GYVGLALAGLAALPLASGEYGASPLCLPYAPPEQAPALGTVLAVMNSCFELVYA 414
 Db 174 IMGWVCCFELLALPLVGVISSTAKVSIQCPM--DTETPLALAVIVFLTLNIVAFYIVC 231
 QY 415 GAVIKLYCDLPBGDFE--AVWDCAMVRHVALIFADGILLYCPVAPLFSASMLGPPVPEA 473
 Db 232 CCHVKIYITVRNPQYNPBGDKDTIARMAVLITDFICMAPISFYALSALINKPLITVSN 291
 QY 474 VASVLLVLPACNLPLLYLLENPHFRDRLRLRPAGDSGLAYAAAGEL--EKSSCD 531
 Db 292 SKILLVLFYPLNSCANPFLYAIPTKAFQRODVETILLSKFGICKRQAOAYRQORVPKKNSTD 351

RESULT 15

US-09-862-767A-5
 ; Sequence 5, Application US/09862767A
 ; Patent No. US20020034786A1

GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCPICN1(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; CURRENT FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: US 09/127,856
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 458
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-862-767A-5

Query Match 9.6%; Score 319; DB 10; Length 458;
 Best Local Similarity 24.9%; Pred. No. 1.4e-17;
 Matches 103; Conservative 77; Mismatches 213; Indels 20; Gaps 8;

Sat Nov 9 13:04:29 2002

us-09-851-595-5.ra1

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:33:15 : Search time 12.9467 seconds
(without alignments)
1438.571 Million cell updates/sec

Title: US-09-851-595-5

Perfect score: 3307
Sequence: 1 NTMYRESWACRYRSIGP.....GGLSGGGGFGPQSLAFASHV 633

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCNUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	Match length	ID	Description
1	958	29.0	644	2	US-08-866-757-2
2	958	29.0	644	4	US-09-153-593-2
3	419.5	12.7	764	4	US-07-741-453A-60
4	419	12.7	696	4	US-07-757-342D-6
5	417	12.6	692	4	US-07-757-342D-6
6	414.5	12.5	695	3	US-08-487-886-2
7	414.5	12.5	695	3	US-08-482-855-2
8	414.5	12.5	695	3	US-08-474-986-2
9	414.5	12.5	764	4	US-07-741-453A-54
10	407.5	12.3	764	4	US-07-757-342D-3
11	407.5	12.3	764	4	US-07-741-453A-61
12	406.5	12.3	764	4	US-07-757-342D-5
13	403.5	12.2	764	4	US-07-741-453A-59
14	381	11.5	674	4	US-07-757-342D-10
15	381	11.5	699	4	US-07-757-342D-2
16	366.5	11.1	792	4	US-07-741-453A-56
17	356	10.8	792	4	US-07-741-453A-55
18	354	10.7	611	4	US-07-757-342D-8
19	354	10.7	636	4	US-07-757-342D-7
20	343.5	10.4	764	4	US-07-741-453A-29
21	318.5	9.6	420	4	US-08-795-876-33
22	318.5	9.6	423	4	US-08-795-876-38
23	318.5	9.6	436	4	US-08-795-876-2
24	313	9.5	336	1	US-08-118-270-54
25	313	9.5	336	1	US-08-118-270-55
26	289.5	8.8	327	5	US-08-118-270-55
27	289.5	8.8	327	5	PCT-US93-08528-55

28	250.5	7.6	332	1	US-08-118-270-53
29	250.5	7.6	332	5	PCT-US93-08528-53
30	140	4.2	673	4	US-09-063-950-2
31	135.5	4.1	370	4	US-08-900-230-2
32	135.5	4.1	370	4	US-08-900-230-2
33	134.5	4.1	369	4	US-09-058-333A-2
34	134.5	4.1	370	4	US-09-172-353-6
35	134.5	4.1	372	2	US-09-199-737-2
36	134.5	4.1	372	2	US-08-626-685A-8
37	134.5	4.1	372	4	US-08-993-088A-2
38	134.5	4.1	372	4	US-08-993-424B-2
39	134.5	4.1	372	4	US-08-665-034A-2
40	134	4.1	371	4	US-08-665-034A-4
41	134	4.1	371	4	US-08-993-424B-9
42	131	4.0	350	2	US-08-466-103A-12
43	131	4.0	353	2	US-08-466-103A-12
44	131	4.0	361	3	US-09-034-985-2
45	130.5	3.9	372	4	US-08-993-088A-20

ALIGNMENTS

RESULT 1
US-08-866-757-2
Sequence 2, Appli
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOURAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGMA, DEK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAAT1)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2
Query Match 29.0%; Score 958; DB 2; Length 644;
Best Local Similarity 36.6%; Pred. No. 7.8e-76;
Matches 211; Conservative 85; Mismatches 191; Indels 60; Gaps 9;

```

OY 54 KLENDTFSQSLQALDLSWMAIRSIHPEARSTLHSLVKLDTNQLTTLPLAGLGIMH 113
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 72 QIKETFGGLISRLDLSRLNLIHEISRATGPIITNLPVSENELTSFTEGLINGLNO 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 114 LKLGKLNALSAFSDSPFKRLILEVPAVOCCEYGMCAFEKASGQWEADLHDD--- 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 132 LKLGKLNALSAFSDSPFKRLILEVPAVOCCEYGMCAFEKASGQWEADLHDD--- 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 171 -EESKRRLGLLARAQENHHYDQDDELQLEMDSKPHPSVOCSPPTGPFKCEYLFEESWG 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 188 AOEKGTADAANVTSTLENEHSOI-----IHCPTSGAFKFCCEYLFGSWM 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 230 IRLAVMAIVLTLVLCNGVLTLTFVAGGAPLPVYKFWGAIAGANTLTGISCGLIASVDA 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 234 IRLTVMFIFLVALFNLVLTITTFASCTSLPSSKLEIGLISVNLPMGIYTGILFLDA 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 290 LTFGPFSEYGARMTGICGRATGFLAVLGSSEAVLTLTAAVOCVSVSCVRAAGKPSL 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 VSMGRFAEFGIMWETGSGCKVGTGLAVFSESSEAFILMLATVERSLSAKIMKNGSNHL 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 350 GSVRAGVIGCLALAGLAAALPLASVGEYASPLCLPYAPPEGCPALGFTVALVMMNSFC 409
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 KOPRVAALAFIATGATVTCGCPFLPHRGEYSASPLCLPF--PTGETPSLGTFTVTLVLSLA 410
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 410 FLVAVAGATIKYCDLPRGDFEAVMDCAMVRHVAWLIFADGLICYCPAFLSFASMLGLFVY 469
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 411 FLMAVITKLYCNLEKEDLSSENSOSSMIKHVAWLIFTCIFCVPAAFFSEAPLITAIISI 470
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 470 TPEAVKSVLLVPLPACLNPLVLTLEPNHFRD---LRRLPRAGDSGLAYAAAGEL 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 471 SPEIMKSVTLIFPLPACLNPLVLTLEPNHFRD---LRRLPRAGDSGLAYAAAGEL 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 526 EK-----SSCDSTQALVAFSVD---LLEASBAGRPPGLETYGFPSVT 566
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 531 EODEFYDDGMSHLOGNLTVCDCESPFLTKRPVSKHLI-----KSHSCPALA 578
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 567 LISCOOP 573
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 579 VASCOOP 585
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 2

```

US-09-153-593-2
; Sequence 2, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIATONG
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAAT1)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153, 593A
; EARLIER FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866, 757
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 644
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-153-593-2

```

Query Match 29.0%; Score 958; DB 4; Length 644;
 Best Local Similarity 38.6%; Pred. No. 78-76;
 Matches 211; Conservative 85; Mismatches 191; Indels 60; Gaps 9;

```

OY 54 KLENDTFSQSLQALDLSWMAIRSIHPEARSTLHSLVKLDTNQLTTLPLAGLGIMH 113
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 72 QIKETFGGLISRLDLSRLNLIHEISRATGPIITNLPVSENELTSFTEGLINGLNO 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 114 LKLGKLNALSAFSDSPFKRLILEVPAVOCCEYGMCAFEKASGQWEADLHDD--- 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 132 LKLGKLNALSAFSDSPFKRLILEVPAVOCCEYGMCAFEKASGQWEADLHDD--- 187
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 171 -EESKRRLGLLARAQENHHYDQDDELQLEMDSKPHPSVOCSPPTGPFKCEYLFEESWG 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 188 AOEKGTADAANVTSTLENEHSOI-----IHCPTSGAFKFCCEYLFGSWM 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 230 IRLAVMAIVLTLVLCNGVLTLTFVAGGAPLPVYKFWGAIAGANTLTGISCGLIASVDA 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 234 IRLTVMFIFLVALFNLVLTITTFASCTSLPSSKLEIGLISVNLPMGIYTGILFLDA 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 290 LTFGPFSEYGARMTGICGRATGFLAVLGSSEAVLTLTAAVOCVSVSCVRAAGKPSL 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 VSMGRFAEFGIMWETGSGCKVGTGLAVFSESSEAFILMLATVERSLSAKIMKNGSNHL 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 350 GSVRAGVIGCLALAGLAAALPLASVGEYASPLCLPYAPPEGCPALGFTVALVMMNSFC 409
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 KOPRVAALAFIATGATVTCGCPFLPHRGEYSASPLCLPF--PTGETPSLGTFTVTLVLSLA 410
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 410 FLVAVAGATIKYCDLPRGDFEAVMDCAMVRHVAWLIFADGLICYCPAFLSFASMLGLFVY 469
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 411 FLMAVITKLYCNLEKEDLSSENSOSSMIKHVAWLIFTCIFCVPAAFFSEAPLITAIISI 470
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 470 TPEAVKSVLLVPLPACLNPLVLTLEPNHFRD---LRRLPRAGDSGLAYAAAGEL 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 471 SPEIMKSVTLIFPLPACLNPLVLTLEPNHFRD---LRRLPRAGDSGLAYAAAGEL 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 526 EK-----SSCDSTQALVAFSVD---LLEASBAGRPPGLETYGFPSVT 566
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 531 EODEFYDDGMSHLOGNLTVCDCESPFLTKRPVSKHLI-----KSHSCPALA 578
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 567 LISCOOP 573
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 579 VASCOOP 585
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 3

```

US-07-741-453A-60
; Sequence 60, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 1991/015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:

```

RESULT 4
 US-07-757-342D-4
 : Sequence 4, Application US/07757342D
 : Patent No. 6218509
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: IGARASHI, Masao
 : MINEGISHI, Takashi
 : NAKAMURA, Kazuo
 :
 : TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 :
 : NUMBER OF SEQUENCES: 10
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
 : CUSHMAN
 :
 : STREET: 130 Water Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: US

```

1      ZIP: 02109
2
3      COMPUTER READABLE FORM:
4
5      MEDIUM TYPE: floppy disk
6
7      COMPUTER: IBM PC compatible
8
9      OPERATING SYSTEM: PC-DOS/MS-DOS
10
11     SOFTWARE: PatentIn Release #1.0, Version #1.25
12
13     CURRENT APPLICATION DATA:
14
15     APPLICATION NUMBER: US/07/757,342D
16
17     FILING DATE: 10-Sep-1991
18
19     CLASSIFICATION: <Unknown>
20
21     ATTORNEY/AGENT INFORMATION:
22
23     NAME: BUCKLEY, Linda M.
24
25     REGISTRATION NUMBER: 31003
26
27     REFERENCE/DOCKET NUMBER: 41226
28
29     TELECOMMUNICATION INFORMATION:
30
31     TELEPHONE: (617)523-3400
32
33     TELEFAX: (617)523-6440
34
35     TELEX: 200291 STRE UR
36
37     INFORMATION FOR SEQ ID NO: 4:
38
39     SEQUENCE CHARACTERISTICS:
40
41     LENGTH: 696 amino acids
42
43     TYPE: amino acid
44
45     TOPOLOGY: linear
46
47     MOLECULE TYPE: protein
48
49     SEQUENCE DESCRIPTION: SEQ ID NO: 4:
50
51     US-07-757-342D-4

```

	Query Match	12.7%	Score 419;	DB 4;	Length 696;
	Best Local Similarity	25.2%;	Pred. 2.4e-28;		
	Matches 134;	Conservative	96;	Mismatches 24;	Indels 56; Gaps 12
QY	26	VERSGQLSPAHPSIALAALASNTFASGKLEXPFTFOLS-----LQALDLSMNA-IR	77		
DB	153	LEICDNHITTVAN-AFGQNMNESTLKLKYGNFELIOSHANGTLLISLEKEVNAHKL	211		
QY	78	SIHPAEFTSLHSVKLDTDNQLTTPPLAGLGLMLKLKGNLALSOAFESKDSFPKRL	137		
DB	212	KMHNDARGARGPSIIDISSTKLOALPSYGLIESIOTLIATSSYLKLPRERFTNLDA	271		
QY	138	EVRYAYOCCTY-----GMCASFASQOMEADHLIDEESSKRPILSILARQAEHNY	189		
DB	272	TLTPSHCCAFRLNPTKEONFSSTFK-----NSKQCESTARRP-----NNETLY	317		
QY	190	-----DQDLDELQLEMEDSKRHPVSQSPPTGPPKPCERYLFESNGITLAWAVALVLSLC	244		
DB	318	SATPAESELSPMDYDQPCSP-KTLQCAPBEDPANPCEDIMGYDFRLIWLINILIMG	376		
QY	245	NGIYLLTVFAGGPAPLDPVKFVCAIAGANTLPISGGLASVALAFEGQSESGARMT	304		
DB	377	NVTVLEVLITSHYKLTVP-RELKCNLSPADFCMGILYLLILIASVDQTKGQYNNIAIDMQT	435		
QY	305	GLGGRATGELALVISEASVLLLTAAVQSCVSVCRVYGRKSPISGSRVAGVGLCLAG	364		
DB	436	GNQGSVAQEFFVPSSELSVTLITVITTLERHHTIYAIDLDQKLRILRHAIPIMLGWMLFST	495		
QY	365	LAALPLIASVGEYCASPICLPYAPPEGPALGFTVALVMNNSPCEITVANAAGYIKLYCDL	424		
DB	496	LIAMLPGLVGSYMSKVSICLPW--DVETTLISQVYLLTLLILINVAFTIICACYIKITFAY	553		
QY	425	PRGDEAY-WDCAMRHVAMLIIFADGLLYCPAFISFASMLLFLPVTFEAYKSVLLVLP	483		
DB	554	ONPELMAINIKDKIKAKNAVLIFTFCTMAPISFPAISALALKVLPILITYNSKVLVLYLP	613		
QY	484	LPACINPLIYLLFNPHFRDDLRLRPRAGDSGLAYAAAGELKESCDSTQA	555		
DB	614	VASCANPFLIYALTFKAFRDFLL-----LSKGGCCKHQA	648		

```

RESULT 5
US-07-757-342D-6
; Sequence 6, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:

```

```

APPLICANT: IGARASHI, Masao
            MINESIGISHI, Takashi
            NAKAMURA, Kazuto
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
            CUSHMAN
            STREET: 130 Water Street
            CITY: Boston
            STATE: Massachusetts
            COUNTRY: US
            ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 692 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-07-757-342D-6
Query Match      12.6%  Score 417;  DB 4;  Length 692;
Best Local Similarity 25.0%  Pred. No. 3.6e-28;
Matches 167;  Conservative 86;  Mismatches 251;  Indels 164;  Gaps 23;

QY 16 SGIPGTHASVRSOGL-----SLPAHPSALALAASTTASGKLEXTDFQSLSL 66
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 92 SNLPKRLHRIEKANLLYINEAFQNP-----SLRYLLISNT-----GKHLPAVHKIQSL 144
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 67 Q-----ADLSWNAIRSIHPEAF-----S 85
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 145 QKVLDDIQDNIHIVANNSFMGISESVIIMLSKNGIEEIHNCADFNGTQDLELNLSDNN 204
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 86 TLHSL-----VKLDLTNDQLTTLPLAGLGGLMLKGNLALSQASKSDFKL 134
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 205 NLEELPNDVFOGASGPVLLDSRTKRVHSLPHNGLENLKKLRARSTYRLKKLPDLKFEVTL 264
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 135 RILEVPAYQCCPRGKASGKQWAEDELH-----LD 169
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 265 MEASLTPSHCCAF-----ANLKRQISELHPICKNSILRODIDMTQIGDQVSLID 316
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 170 DESSSKRPLGLARQAEHHYDQDELQLEMEDESKPHPSYQCSPTPGPKCEYLFESNG 229
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 317 DEPS-----YKGSMDMNNEDPDYDCN-----EYVD-----VTCSPKPDAFNCECDIMGYNI 363
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 230 IRLAWNAIYLVLCNGVLLLVFAGGAPAPLPYKFFVGAIGAGANTLTGISCGLIASYDA 289
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 364 LRVLIWFISILAITGNTTVLV-VLTTSQYKLVPRFLMNCNIAFADLCIGIYLLIASYDI 422
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 290 LTFGQSEYGARRETEGLGATGFLAVLGSEASYLLTLAAVQ-----CSVSVCYRAY 343
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 423 HTKSQYHNVAIDMOTAGCGAAGFTVFASLSVYTLTATLEWHHTITHAMQLECC----- 478
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 344 GSPSLGAVRAGVLCIALAGLAALPLASVGEYGASPLDLPVAPPGCGAPALGFYALV 403
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

Db 479 -KVQLRHAASVMVLG-WTFAPAAALFIFGISTYKWSICLPM--DIISPLSQLYVMALL 534
QY 404 MNNSFCFLVYAGATIKLYCULPRGDF-EAYVDCAMVRHVAMLIADGLLYCPVAFISFAS 462
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 535 VLVNLAFFVVGCGTYHLYLVNRPNTIVSSSDTIKAKRMATLIFTDFLCMPISFFALSA 594
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 463 MGLGFVTPPAVKSVLVLPPLPACLPILYLFNPHPRDRLRLRRAGDSGLPATAAA 522
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 595 SLKVPITTVSKAKILLVLFYPINSCANPFLAITTKNFRDFTL----- 639
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 523 GELEKSSCDSTQALVAFSDVDLLIEASEAGRPGLTTFYGPSPVYLISCOQGPARELSH 582
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 640 --LSKFCGYEMOAOIYRT-----ETSSA-----THNF-HARKSOS--SAPRVVNSY 681
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 583 CVPDEGNH 590
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 682 VLVPL-NH 688

RESULT 6
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 5744448
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 5744448een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Stephan P. Williams,
ADDRESS: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massil via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular

```


IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2

Query Match 12.5% Score 414.5; DB 1; Length 695;
Best Local Similarity 25.0%; Pred. No. 6e-28;
Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;

55 LENDPFSOLS-SLQALDLSWNAIRSIHPEAF-----STLHSL----- 90
160 IERNVSGVLSFESVILMLNKGIOETINCAFNQDLELNSDNNNEELPNDVPHGASG 219
91 -VKLDLTQDQTLTPLAGLGLMLHLKGNLAISQAFSKSPFLRIIEVPAYOCPPYG 149
220 PYLIDISRTIRHISPGSGLNKLKIRASTYNNLKLPTLEKLVAMEASTLYPSHCAF- 278
150 MCASFPAKSGOWAEPLH-----LDDESSKRPGLLARQAEHNDYDQDDELQJL 198
279 -----AMRRQISELHPICKNSILRQEVDMYTQTRGSSSLAEDNESSYSRQFDMYTT 331
199 EMDSPKHP--SVQCSPTPQPFKPCCEYLFEWSGIRLAWALVLLSVLCNGVLLTFVAGG 256
332 EPDYDLCLNEVVDYTCSPKPAFNPCEIDIMGVNIILRVLIWFSILAIIGN-IIIVVILITS 390

257 PAPLPKFEVGAIGANTLTIGISGLASVDALFFGFSYGARETGLCRATGFLAV 316
391 QYKLVPRFLMGNLAFADICIGIYLLIASVDIHFTSQXHNVAIDQTSAGCDAAEFYV 450
317 LGSEASVLLITLAAVQ-----CSVSVQVRAVGRKPSISGVSAGVIGCLALAGLAAAP 370
451 FASELSVYTLTAITLERMHTITHAMQDC-----KVLRHNASVWVGWT-EFAAALFP 504
371 LASVEYGASPLCLPYAPEGPPAALAGFTVALVMMSFCELVVAGAYIKLYCDLPBGDF- 429
505 IFGISYMKVSTCLPM--DIDSPLSQLYMSLLVNLAVLAFVVICGXIHYLLVRNPV 562
430 EAWDCAMVRHVAWLIFADGLICPVAPLSPASMLGFLPYTPRAVSVLLVLPACLN 469
563 SSSSTRLAKRMAMLIFTDFLCMAPISFAISASLSKVLPLTVSKAKILLVLFHPINSCAN 622
490 PLLYLFPNFERDDLRLRPRAGDSGLAVYAAAGELEKSDCST 533
623 PLVIAIFTKNFRDFILSKCG-----CYEMOQILYRTETST 661

RESULT 7
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016en Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massil via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482, 855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670, 085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain

```

? LOCATION: 1 to 349
? IDENTIFICATION METHOD: similarity with other
? IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
? IDENTIFICATION METHOD: domains, hydrophilic
? FEATURE:
? NAME/KEY: transmembrane domain
? LOCATION: 350 to 613
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
? FEATURE:
? NAME/KEY: putative transmembrane region I
? LOCATION: 350 to 370
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative transmembrane region II
? LOCATION: 382 to 404
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative transmembrane region III
? LOCATION: 427 to 448
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative transmembrane region IV
? LOCATION: 469 to 491
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative transmembrane region V
? LOCATION: 512 to 533
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative transmembrane region VI
? LOCATION: 557 to 580
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative transmembrane region VII
? LOCATION: 592 to 613
? IDENTIFICATION METHOD: similarity to other G
? IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
? IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
? FEATURE:
? NAME/KEY: putative carboxy-terminal intracellular
? NAME/KEY: domain
? LOCATION: 614 to 678
? US-08-482-855-2

Query Match 12.5%; Score 414.5; DB 3; Length 695;
Best local Similarity 25.0%; Pred. No. 6e-28;
Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;

? 55 LEXDPTFQSLS-SLQALDLSWMAIRSHPEAF-----STLHSL----- 90
? 160 IERNSEFVGLSFESEVILMLNKNGIOEIHNCACFNGTQDELNLSDNNLLELPNDVFGASG 219
? 91 -VKLDLDNLDTITPLAGLGLMLHLKGNLALSAFSDSPFKRLITLVPAYQCCPYG 149
? 220 PVLDISRTIHSHPSYGLKLEKRLARSTYLNKLPLEKLVALMEASILTYPSHCAF- 278
? 150 MCASFASQGWMEADLH-----LDDESSKRPGLGLARQAEHNYDDLDLEQL 198
? 279 -----ANRRQISELHPICNKSILRQEVVYMTOTRGORSSLAEDNESSYSGFMTYT 331

```

```

? 199 EMEDESKPH--SVQSPPTPGPKPCXEYLFESWIGIRLAWAIVLVLSCNGVLTTFVAGG 256
? 332 EFDYDLCNEVVDYTCSPKPAFNPCEIDIMGYNILRVLIWFIISITATGN-ITVYILTS 390
? 257 PAPLPVKRFVGAIAQANTLTGISGILLASVDALTFQSEYGARWETGLACRATGFLAY 316
? 391 QXKLTVPRFLMCLNLAADLCIGYLLLIASVDITHKSOYHNVAIDWQTAGCADAAGFFTV 450
? 317 LGSEASVLLTLAAV-----CSVSYCYRAVKSSLSGVRAGVIGCLALAGIAALP 370
? 451 FASELSVYTLTALTTERWHTITHAMQDLC---KVQLRHAASVVMGMW-FAPFAALFP 504
? 371 LASVGEYASPLCLPAPPEGPAALGFVALVYMNSECFVLVAGAYKLYCDLPBGDF- 429
? 505 IGCISYMYVSLCPM--DIDSPLSQLYVMSLVLVNLAVFYIGCYIHITLVANPNIV 562
? 430 EAVWDCAVRHVAWLIFADGLLYCPVAFLSPASMLGLEPVTPEAVKSVLVLPLPACLN 489
? 563 SSSDPRIAKRMAMLIPTDFLCMAPISEFAISALIKVPLITVSKAKILLVLFHPINSCAN 622
? 490 PLLYLFFNHFRRDLRLRRADSGPLAFAAGELKSSCDST 533
? 623 PFLYALFTKFNFRDPFILSKG----CYEMQAQIVRTTST 661
? DB

```

```

? RESULT 8
? US-08-474-986-2
? GENERAL INFORMATION:
? APPLICANT: Kelson, Christie Ann
? Schweißhardt, Rene Lynn
? Cheng, Shirley Vui Yen
? Nugent, No. 6372711een Patrice
? TITLE OF INVENTION: Human follicle stimulating
? HORMONE RECEPTOR
? NUMBER OF SEQUENCES: 2
? CORRESPONDENCE ADDRESS:
? ADDRESSSEE: Stephan P. Williams,
? Ares-Serono, Inc.
? STREET: Exchange Place, 37th floor
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
? OPERATING SYSTEM: MS-DOS version 4.0
? SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/474, 986
? FILING DATE: 07-Jun-1995
? CLASSIFICATION: <unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/670, 085
? FILING DATE: <unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Williams, Stephan P.
? REGISTRATION NUMBER: 28546
? REFERENCE/DOCKET NUMBER: US/252
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 723-1300
? TELEFAX: (617) 723-8923
? LOCATION: 614 to 678
? SEQUENCE DESCRIPTION: SEQ ID NO: 2:
? US-08-474-986-2

Query Match 12.5%; Score 414.5; DB 4; Length 695;
Best local Similarity 25.0%; Pred. No. 6e-28;
Matches 131; Conservative 94; Mismatches 232; Indels 67; Gaps 13;

? 55 LEXDPTFQSLS-SLQALDLSWMAIRSHPEAF-----STLHSL----- 90
? 160 IERNSEFVGLSFESEVILMLNKNGIOEIHNCACFNGTQDELNLSDNNLLELPNDVFGASG 219

```

```

OY 91 -VKLDLTNOQLTTLPLAGIGLMHLKGNLALSOAFKSDSPFKRLILEVPAYOCPPY 149
OY 220 PYLIDISRTIRHSLSBYGJENLKLARSTYLNKLPLEKLVALMESASTYPSHCAF- 278
OY 150 MCASFPGASGQWAEHL-----LDDESSKRLGLLARQAEHNYDODLDELQ 198
OY 279 -----AMNRROISLHPICKNSILROEVDYWTOTRGORSSIAEDNESSYSGFDMTYT 331
OY 199 EMEDSKPH--SYGSPPTGPFKPCYELFESWIGRLAWAIVLISVLCNGVILTYFAGG 256
OY 332 EFDYDLCEVVDYTCSPKPAFNPCEIDMGXNLRVLIWFISLATTGN--ILVLTITS 390
OY 257 PAPLPVKFVAGIAGANTLTGISCGLASVADLTGOFSEYGARMTGLGCRATFLAY 316
OY 391 QYKLTVPREFLMCNIAFADLCIGYLLILASVDLHTKSQYHNRYAIDMOTGACGCAAGFTV 450
OY 317 LGSFASVILLFLAVO-----CSVSVCYRAXGKSPSLGSRAGVIGCLALAGLAALP 370
OY 451 FASLSTYTLTALTLERHNTTHAMQDLC-----KVALRHAAVYVWGWI--FAFAALFP 504
OY 371 LASVGEYASPLCLPYAPPEGQAPALGFTVALMNSFCFLVAGAYIKLYCDLPBGDF- 429
OY 505 IFGISSYKVCISLPM--DIDPSLQILYMSLVNLVLAIFYVIGCYTHIYLVTRNPIYV 562
OY 430 EAVWDCAMVRAHVAFLIFADGLIYCPAFSLFASMLGLFPVTPPAVKSVLVLPPLACLN 489
OY 563 SSSSTRIKRAKMAILTFDFLCMAPISTFPAISASLKVPLITYSKAKILLVPHPLINSCAN 622
OY 490 PLIYLLFNHFRDRLRLRPFRAGDSGLPAAAGLEKSSCOST 533
OY 623 PLYAIFTKNFRDFEILLKCG-----CYEQAOITYRRETST 661
OY 623 PLYAIFTKNFRDFEILLKCG-----CYEQAOITYRRETST 661

RESULT 9
US-07-741-453A-54
; Sequence 54, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEO ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 amino acids

```

```

; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-741-453A-54

```

```

Query Match 12.5%; Score 414.5; DB 4; Length 764;
Best Local Similarity 23.7%; Pred. No. 6,9e-28;
Matches 141; Conservative 102; Mismatches 258; Indels 95; Gaps 12;

```

```

OY 33 SUPAHASIALAASNTASCKLENDPESOL-----SSIQALDLISWNA--IRSIHPEAF 84
OY 166 SIPAN-----AFQGLCNEVTLTKLYNKGFTSTQGHAFNFTKIDAVLYNKNKYLSDADAF 221
OY 85 STLHS-LVKLDLTNOQLTTLPLAGIGLMHLKGNLALSOAFKSDSPFKRLILEVPAY 143
OY 222 GGYISGPTLLDDSYSTVALPFGKLEHKEIYARTWTWTLKLPISLSLFLHTRADLSIPS 281
OY 144 OCCPY-----GMCSF-----FKASG 159
OY 282 HCCAFNRNKKIRGILLESIMCNESSTLSLRKSVNTLNGPFOEYEEYLGDSHAGYKDNS 341
OY 160 QW-----EADLHLDESSKRLPLGLLARQAEHNYDODLDELQEMEDSKP 205
OY 342 QFQDDBNSHYVFFEEQDEILFGQELKNPOETLQAFDSHYDTV-----CGG 392
OY 206 HPSYQSPPTGPFKPCYELFESWIGRLAWAIVLISVLCNGVILTYFAGGAPLPVKF 265
OY 393 NEDVVCYCPKDEFPKCDIMGKFLRYVWFVSLALGVFVILYLTSHYKLYP--RF 451
OY 266 VVAGIAGANTLTGISCGLASVADLTGOFSEYGARMTGLGCRATFLAVLSEASVLL 325
OY 452 LMCNIAFADFCMGYLLILASVADLTHTSEYVNHAIWMOTGGCTAGFTFYFASLSVYT 511
OY 326 LTLAAVQCSVSVCYRAXGKSPSLGSRAGVIGCLALAGLAALPLASVGEYASPLCLP 365
OY 512 LTVITLERMYALTTFAMRLDKRIHLRAHAIWVGWCCFLALLPLVGISSYAVSICLP 571
OY 386 YAPPEGQAPALGFTVALMNSFCFLVAGAYIKLYCDLPBGDF- AWDCAHVAHVA 444
OY 572 M--DIETPLALAYIILVLLINVAIFIIVSCYKITYITVRNPOYVPGKDKIKRAHVA 629
OY 445 IPADGLIYCPAFSLFASMLGLFPVTPPAVKSVLVLPPLACLNPLLYLNFHFRDRL 504
OY 630 IFTDFCMAPISIFYALSAHMKPLITYNKSILLVLFVPLNSCANPFLYAITKAFQDV 689
OY 505 RRLRPAGDSGLPAAAG-----ELEKSSCOSTQALVAFSDVDILEAS 549
OY 690 FILLSKFGICKRQAQAYRGQVSPKNSAGIOIQKVTYRDMROSIPNNQDEYELLENS 745

RESULT 10
US-07-757-342D-3
; Sequence 3, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; APPLICANT: MINOJISHI, Takashi
; APPLICANT: NAKAMURA, Kazuo
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/757,342D
 FILING DATE: 10-Sep-1991
 CLASSIFICATION: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BUCKLEY, Linda M.
 REGISTRATION NUMBER: 31003
 REFERENCE/DOCKET NUMBER: 41226
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)523-3400
 TELEFAX: (617)523-6440
 TELEX: 200291 STRE UR
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 700 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-07-757-342D-3

Query Match

Best Local Similarity 12.3%; Score 407.5; DB 4; Length 700;
 Matches 135; Conservative 99; Mismatches 254; Indels 75; Gaps 13;

QY 47 SNTTASGKLEXPDSQSS-----LQALDLSWNA-INSIHPEAFSTLSVTKDLDTN 98
 DB 177 NNESTVTKLYGFEVSHANGNTLLISLEKENYLEKMHSGAVGATGSIIDISST 236
 QY 99 QLTTPPLAGLGLMLHKLKGNLALSOAFSKDSFPKRLIEVYVAVOCCPYGMCASFRRAS 158
 DB 237 KLOALPSSHLSIQTLILASSYSLTKLPSEKFTSLVATLVTPSHCCAFRLPK----- 291
 QY 159 GQWEAEDHLDEESSKRPGLLAROENH--YDODDELQIEMED-----SKRPSVQC 211
 DB 292 ---KEONFSFIENFSCQCESTYRKADNETLYSAIFEENELSGWDYDYGCSF-KTIQC 347
 QY 212 SPTGPRPCPYLEESWGIRLAVAVLTVLVCNGVLTVFAGGPARLPYKPVYGAIA 271
 DB 348 APEPDARNPCEIDINGYAFRLVILWILNITLIFGN-LTVFLVLLSRKLYTPRLMCMIS 406
 QY 272 GANTLTGISCGLASVALTFGQSEFGARWETGLCRATGFLAVLGSEASVLLTLTAAV 331
 DB 407 FADFCMGLYLLLASVDSOTKGOYNNHAIIDWQSGCGAAGFFTVFASSELVYTLTVTL 466
 QY 332 QCSVSVCAVAGKSPSLGVRAGVLCGLAAGLAALPLASVGEYGASPLCLPYARPEG 391
 DB 467 ERNHTTYAVQIDQKRLRLRAIPIMLGGWLFSTLIATMPLVGISNMYKVSICLPM--DVE 524
 QY 392 QPAALGFVALVMNSFCFLVAVGATIKLDCDLPRGDFEAV-WDCAMVRHVAWLIFADGL 450
 DB 525 STLSQYVILITLNVAVVICACIIRITFAVONDELTAAPNKDTRKIAKMAILITFDFT 584
 QY 451 LVCPAFLSFAASMLGLPYTPPEAVKSVLLVLPACILNPLYLLNPHRDLRLRLR 510
 DB 585 CMAPISFEFAISAFKPLVTLVNSKILLVFPVNSCANPFLAITAKAFORFELLILSR 644
 QY 511 AGDSGPLVAAAGLEKSCDSTQALVAFSDVDLIIEASGAPRGLEY-----GFP- 563
 DB 645 FG-----CCKRA-----ELVRRKEFSAYTSNCKNGPFG 673
 QY 564 -----SVTLISQCGPAGPR 577
 DB 674 ASKPSQATLKLSTVHCOPIPPR 696

RESULT 11
 US-07-741-453A-61

Sequence 61 Application US/07/41453A
 Patent No. 6228597

GENERAL INFORMATION:
 APPLICANT: PARMENTIER, MARC

APPLICANT: LIBERT, FREDERIC
 APPLICANT: DUMONT, JACQUES
 APPLICANT: VASSART, GILBERT
 TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
 TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN DAREY & CUSHMAN
 STREET: 1615 L STREET, N.W.
 CITY: WASHINGTON, D.C.
 COUNTRY: U.S.A.
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/741,453A
 FILING DATE: 19911015
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KOKULIS, PAUL N.
 REGISTRATION NUMBER: 16773
 REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 861-3000
 TELEFAX: (202) 822-0944
 TELEX: 6714627 CUSH
 INFORMATION FOR SEQ ID NO: 61:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 764 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-741-453A-61

Query Match

Best Local Similarity 12.3%; Score 407.5; DB 4; Length 764;
 Matches 145; Conservative 99; Mismatches 261; Indels 105; Gaps 13;

QY 19 PGSTHASVRSQGLSLPAHPRASIALAASNTTASGKLEXPDSQSS-----SSIQALDL 71
 DB 162 PYMTSIPVNAFQGL-----CNETLTCLKYNGFTSVQCYAFNGTKLDIAVYL 207
 QY 72 SWNA-INSIHPEAFSTLSLVK-LDLTDNQLTTPPLAGLGLMLHKLKGNLALSOAFSKD 129
 DB 208 NKKKYLTVIDKDAFGVGYSGPSLIDVQTSVTALPSKGLHKLKELIRANNTWTLKPLSL 267
 QY 130 SFPKRLIEVYVAVOCCPY-----GMCSF----- 154
 DB 268 SFLHTRADLSTYSHCCAFRNQKIRGILLESIMCNESSMOSLRQKRSVNALNPLHQEYE 327
 QY 155 -----FKASGV-----EADLHLDEESSKRPGLGLARAEHNYDQ 191
 DB 328 ENLGDSTVIGYKRSKPODTHNNAHYVFPFEDDELIGGQELANKQDETLQAFDSHYD 387
 QY 192 DDELQIEMEDSKRPHSVYCCSPTPGPFKCEYLFESWGIRLAVMAVLLTVLVCNGVLTLT 251
 DB 388 TL-----CGDSE---DWCTPKSDPFNCEIDIMGYKFLRIYVWVSLLALGNFVLL- 437
 QY 252 VFAGGPAPLPYKPVYVGAIGANTLTGISCGLASVDALTFGQSEFGARWETGLCRAT 311
 DB 438 ILLTSHYKLNVPRLMCMIAFADFCMGWILLIASVDLYTHSEYNNHAIIDWQSGCWTA 497
 QY 312 GLFVAVGSEASVLLTLTAAVOCSVSVSCVRAVYKSPSLGVRAGVLCGLAAGLAALPL 371
 DB 498 GFTTVFASSELVYTLTVTLTTERWYAITFAMRLDRIRLRHAAAIWVGWCCFLALPL 557
 QY 372 ASVGEYASPLCLPYPAPPEGOPALGFVALVMNSFCFLVAVAGATIKLDCDLPRGDFE- 430

Db 558 VGISSYAKVSICLPM--DTEPLALAYIMSVLNI VAFVIVCCYKITYTVRNPQNP 615

QY 431 AVMOCAMRVHVAWLIFADLLYCVAFLSFASMLGFPYTPRAVKSVLVLPPLACLP 490

Db 616 GDKDTKIKRAKAVLIFDFTICMAPISFVLSAILNKPLITVSNKILLVLFYPLNSCAMP 675

QY 491 LLYLLEPHFRDRLRLPRAGDSGLPAYAAG-----ELEKSSCDSQALVAF 539

Db 676 FLVAFITKAFQRDVFIILSKFGICKROAQAYRGQVPRKNSDIDIOVKYTHDMRGGLHNM 735

QY 540 SDVDLLEAS 549

Db 736 EDVYELLENS 745

RESULT 12

US-07-757-342D-5

Sequence 5, Application US/07757342D

Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINAGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES:

ADDRESSEE: DAVID G. CUSHMAN, DIKE, BRONSTEIN, ROBERTS & CUSHMAN

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: HUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 764 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-07-757-342D-5

Query Match 12.3%; Score 406.5; DB 4; Length 764;

Best Local Similarity 23.8%; Pred. No. 3.5e-27;

Matches 145; Conservative 98; Mismatches 262; Indels 105; Gaps 13;

QY 19 PGSTHASVRSQGLSLPAHPASIALAALASNTTASGKLENDTFQSOL-----SILQALDL 71

Db 162 PYMTSLPVNAFOGL-----CNETLTKLYNNGFTSVQGYAFNGTKLDAVYL 207

QY 72 SWNA-IRSHIPEARSTLHSLVK-LDLTDNQTITPLAGLIGIMHLKLGKGNALASQAKSD 129

Db 208 NKNKYLTVIYKDAFGVSGPSLSDVSTQSTVTAALPSKGLHLKELIARNWTTLKKPLSL 267

QY 130 SFPRKRLLEVPYAVQCCPY-----GMCAFSF----- 154

Db 268 SFHLHTRADLSYPSHCCAFERKQKIRGILIESLMCNSSMOSLRQKSVNAINSLPHQYE 327

QY 155 -----FKASGQW-----EAEDLHLDDESSKRPLGLLAROENHHYQ 191

Db 328 ENIGDSIVGKREKSRQDHNHNAHYVFFEEQDEITIGQELKNQDEITLQAFSDHYD 387

QY 192 DLDELQLEMEDSKPHPSVQCSPTTPGPKCEYLFESMGIRLAVWAVILSVLNGVLLT 251

Db 388 TI-----CGDSE---DMWCTPKPSDEFPCEIDIMGYKFLNIVWVYSLALLGNVFL- 437

QY 252 VFAGCAPLEPPYKVVGAIGANTLTGISGLASVDALTFGQFSEYGARWETGLGCRAT 311

Db 438 ILTSHYKLVNPRFLMGNLAFADRCMKMYLLILASVDLYTHSEYNNHAIIDMOTPGCNTA 497

QY 312 GFLAVLGESEASVLLTLTAAVQSVSVSCVRAVKGKPSLGSYRAGVIGCLALAGLAALPL 371

Db 498 GFTTFVASELSVLTITLTERWYAITFAMRLBKRMLRHACALMGVGCFFLALPL 557

QY 372 ASVGEYASPLCLPYAPPEQPALGFTVALVMNSFCFLVAVAGATIKYCDLPRGDFE- 430

Db 558 VGISSYAKVSICLPM--DTEPLALAYIVFVLTINIYAFVIVCCYKITYTVRNPQNP 615

QY 431 AVMOCAMRVHVAWLIFADLLYCVAFLSFASMLGFPYTPRAVKSVLVLPPLACLP 490

Db 616 GDKDTKIKRAKAVLIFDFTICMAPISFVLSAILNKPLITVSNKILLVLFYPLNSCAMP 675

QY 491 LLYLLEPHFRDRLRLPRAGDSGLPAYAAG-----ELEKSSCDSQALVAF 539

Db 676 FLVAFITKAFQRDVFIILSKFGICKROAQAYRGQVPRKNSDIDIOVKYTHDMRGGLHNM 735

QY 540 SDVDLLEAS 549

Db 736 EDVYELLENS 745

RESULT 13

US-07-741-453A-59

Sequence 59, Application US/07741453A

Patent No. 6228597

GENERAL INFORMATION:

APPLICANT: PARMENTIER, MARC

LIBERT, FREDERIC

APPLICANT: DUMONT, JACQUES

APPLICANT: VASSART, GILBERT

TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS

TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN

STREET: 1615 L STREET, N.W.

CITY: WASHINGTON, D.C.

COUNTRY: U.S.A.

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/741,453A

FILING DATE: 19911015

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOKULIS, PAUL N.

REGISTRATION NUMBER: 16773

REFERENCE/DOCKET NUMBER: 91913/1107/US/ST

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 861-3000

TELEFAX: (202) 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 764 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-741-453A-59

Query Match 12.2%; Score 403.5; DB 4; Length 764;
 Best Local Similarity 23.9%; Pred. No. 6,4e-27;
 Matches 146; Conservative 96; Mismatches 263; Indels 105; Gaps 13;

19 PGSTHASERSOGSLIPRAHPSALALASNTTASGLKEXDFSOI-----SSIQALDL 71
 Db PYMTSIVVNAFOGL-----CNETLTLLKLVNMGFTSVGYAFNGTKIDAVYL 207
 QY 72 SWNA-IRSHPEAFSTLSLVK-LDLFDNOLTTPLAGLGIMHLKLGKGNLALSOAFSD 129
 Db 208 NKNKYLIVIDKDAFGVYSGPSLDVSGTSTALPFSKGLHKLKLIARNTWTKRLPLSL 267
 QY 130 SFPKRLLEPVAYQCCPY-----GMCASF----- 154
 Db 268 SFHLIRADLSYSHCCAFKNKKRIGLIESLMCNESMOSLRKRSVNALNSPLHQEYE 327
 QY 155 -----FKASGOW-----EAPDLHDDDESSKRPGLLARQAEHNHDO 191
 Db 328 ENLGSIVGKEKSKPDQTHNNAHYVFFEEQEDIEIGGOELNKPQRETIQAFDSHYD 387
 QY 192 DDELDELMDSDSPHSVOCSPTPGPFCEYLEFESWGRILAWAIVLVLCNGVLTLT 251
 Db 388 T-----CGDSE---DMVCTPKRSDENPCEDINGYKFLRIVVFSLLALGVFVLL- 437
 QY 252 VRAGGAPAPLPVYVGAIGANTLTGISGLASVDALTFGQSESGARWETGLCRAT 311
 Db 438 ILTSHYKLNVPRLMCLNLAFAFCMGMYLLIASVDLYTHSEYNAIDMOTGPGNTA 497
 QY 312 GFLAVIGSASVLLTLAVQCSVSCVAVYKSPSLGSRAGVGLCLAGLAAALPL 371
 Db 498 GFTTFASLSTVYTLVTLERWATTFAMRLDRKIRLRHACALMGWGWCCEFLALLPL 557
 QY 372 ASVGEYASPLCLPYAPRPGOPALGFTVALVMNSFCPLVAGATIKLYCDLPRGFE- 430
 Db 558 VGISYAVVSIOLPM--DETPLALAYIVFVLNLVAIVYVCCCVYKITYITRNPOYNP 615
 QY 431 AVMDCAMRHYAMLIIFADGLLCVPAFLSPASMLGLEPYTPAEVKSVLVLLVPLACLNP 490
 Db 616 GDKDKTIKKRMAVLIETFTICMAPISFYALSAIUNKPLTVNSKILLVLFYLNSCANP 675
 QY 491 LLYLLENFHFRDRLRLRPBGDSGLAVYAAAG-----ELEKSSCDSTQALVAF 539
 Db 676 FLVAIFTRAFORDVFILSKFGICRQAQAVRGQVRPPKNSSTIDIOVKYTHDMRGGLHNM 735
 QY 540 SDVDLILEAS 549
 Db 736 EDVYELIENS 745

RESULT 14
 US-07-757-342D-10
 Sequence 10, Application US/0757342D
 Patent No. 6218509

GENERAL INFORMATION:
 APPLICANT: IGARASHI, Masao
 MINEGISHI, Takashi
 NAKAMURA, Kazuo
 TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
 CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US

ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/757,342D
 FILING DATE: 10-Sep-1991
 CLASSIFICATION: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BOCKLEY, Linda M.
 REGISTRATION NUMBER: 31003
 REFERENCE/DOCKET NUMBER: 41226
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)523-3400
 TELEFAX: (617)523-6440
 TELEX: 200291 STRE UR
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 764 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 US-07-757-342D-10

Query Match 11.5%; Score 381; DB 4; Length 674;
 Best Local Similarity 25.3%; Pred. No. 5.1e-25;
 Matches 122; Conservative 91; Mismatches 241; Indels 28; Gaps 10;

47 SNTTASGKLEXPDTFSGLS-----LQALDSWNA-IRSHPEAFSTLSLVKDLDTN 98
 Db 148 NNESTYTLKLGCFEEVQSHANGTTLTSLKENVHLEKMHNGARGATGPKTDISS 207
 QY 99 QITPLPLAGGLMHLKLGKGNLALSOAFSDSPKRLILEVYAYQCCPYGMCASFEXAS 158
 Db 208 KLOALPSYGLSEIQRLIATSSYSLKKLPBSREFVNLLENTLYPSHCCA-----FRNL 260
 QY 159 GOWEAEHLHDDDESSKRPGLLARQAEHN--YDDDELDELMDSDK-----PHPSVOC 211
 Db 261 PTRKQNSHSHISENFSKO-CGSTVAKRKNKTLYSMLSESLSGMDYEGFCLP-KTPRC 318
 QY 212 SPTGPKPCPEYLEFESWGRILAWAIVLVLCNGVLTLTVFAGGAPAPLPKVFYVGAIA 271
 Db 319 APEPDARFNPCEIDINGYDELRLVLMINILALMGN-MYLFVLLTSRYKLTVPRLMCLNS 377
 QY 272 GANTLTGISGLASVDALTFGQSESGARWETGLCRATGFLAVIGSEASVLLTLAAV 331
 Db 378 FADFCMLYLLILLASVDSQFKQYVNHAIQWOTGSGCSTAGFTTFVASELSYTLTVTL 437
 QY 332 QCSVSVSVAVYKSPSLGSRAGVGLCLAGLAAALPLASVGEYASPLCLPYAPREG 391
 Db 438 ERWHTITYAIIHDOQLRLRAIILMLGWLFSLLIAMLPLVGSNMYKVSICFPW--DVE 495
 QY 392 QPAAIGFTVALVMNSFCPLVAGATIKLYCDLPRGDEAV-WDCAMARHYAMLIIFADGL 450
 Db 496 TLLSOYVITLILNLVNAFFIICACYIKIFYAVRNPELMATNKTAKKAAIILFDFT 555
 QY 451 LYCPVAFSPASMLGLEPYTPAEVKSVLVLLVPLACLNLPLYLENPHFRDRLRLRPR 510
 Db 556 CMAPISFFAISAFAKVLTVINSKVLVLYFPINSCANPFLAIFKTRQDRDFLLSK 615
 QY 511 AG 512
 Db 616 FG 617

RESULT 15
 US-07-757-342D-2
 Sequence 2, Application US/0757342D
 Patent No. 6218509
 GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao
MINEGISHI, Takashi
NAKAMURA, Kazuo
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
CUSHMAN

STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,342D
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BUCKLEY, Linda M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 41226
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR

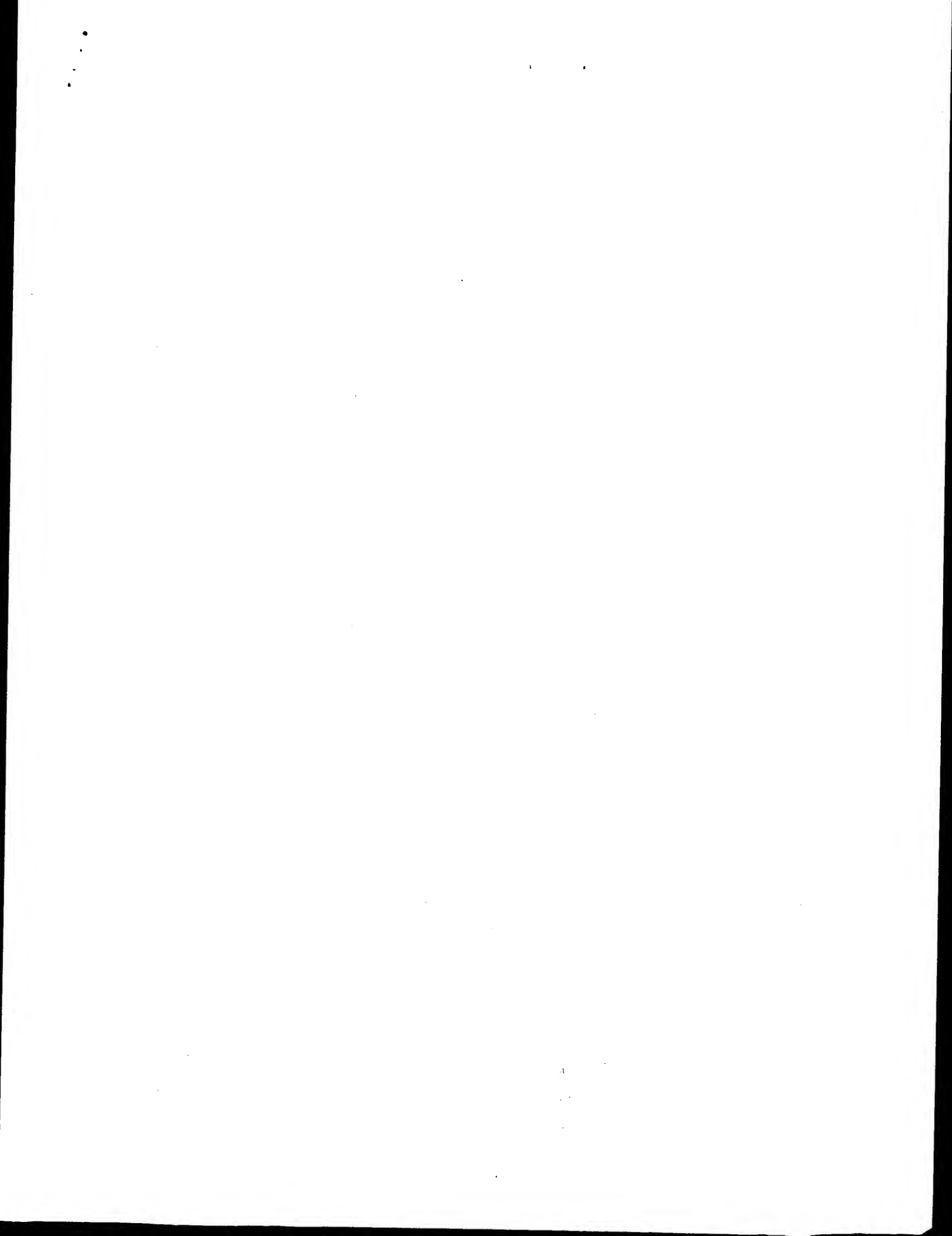
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-07-757-342D-2

Query Match 11.5%; Score 381; DB 4; Length 699;
Best Local Similarity 25.3%; Pred. No. 5,4e-25;
Matches 122; Conservative 91; Mismatches 241; Indels 28; Gaps 10;

Db 581 CMAPISFAISAFAKVPILTVNSKVLVLEYPINSCANPFLIYFTKTFORDFFLLSK 640
QY 511 AG 512
Db 641 FG 642
Search completed: November 8, 2002, 19:37:23
Job time: 16.9467 secs

QY 47 SNTTASGKLENDTFSQSLSS-----LQALDLSMNA-IRSIHPEAFSTLHSLVKLIDLTDN 98
Db 173 NNEGVTLKLVNGEVEVQSHAFNGTTLTSLKENVHLEKMHNGAFRGATGPKTLDISST 232
QY 99 QLTTLPLAGLGLMLKLGKGLALSQAQFSKDSFPKRLILEVPYAVQCCPYGMCASFRRAS 158
Db 233 KLOALPSYGLSTIQLTATSSYSLSKLPSTRETFVNLLEATILTPSHCCA-----FNNL 285
QY 159 GQWEAEDLHLDDESSKRPGLILARQENH--YDQDLDELQLEMEDSK-----PHPSVQC 211
Db 286 PTKQONFHSISENSKQ-CESTYRKVSNKTLYSMLAESELSGWDYEGCLP-KTPRC 343
QY 212 SPTGPPFPCEYLPESMGIRLAVNAIVLLSVLCNGVLVLTVPAGGPAPLPYKFFVGAIA 271
Db 344 APEPDAFNPCEIDIMGYDLRLVILMLITLALGN-WTVLFPVLLTSRKILTVPRFLMCNLS 402
QY 272 GANTLTGISCGLLASVDALTFQFSEYGARWETGLGCRATGFLAVLGSFASVLLTLAAV 331
Db 403 FADFCMGILYLLIAVDSQTKGYNNHAIIDMOTGSGCSTAGFFTFASFELSYTLTVITL 462
QY 332 QCSVSVSCVRAVGKSPSGSVRAGVGLGCLALAGLAAALPLASVGEASPLCLPYAPREG 391
Db 463 ERWHITTYAIHLDOKLRHAILIMLGMFLSSLIAMDPLVGVSNMYKVSICPPM--DYE 520
QY 392 QPAALGFVALVMMNSFCFLVAVGAYIKLYCDLPRGDFEAV-WDCAMVRHVAWLIIFADGL 450
Db 521 TLLSQVYITLITLINVAFIICACYIKIYPAVRNPELMATNMDTKIAKKMALLIFTFDFT 580
QY 451 LVCPAVAFISFASMLGLFPVTPPEAKSVLLVLPACLNPLLYLLFPNPHRDDRLRLRPR 510



PI Gu W;
 XX WPI: 2002-055584/07.
 DR N-PSDB: AA167922, AA167923.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 PT disorder

XX Example 1; Fig 5; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular process, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypothyroidism, hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and pharmacogenomics). The present sequence
 CC represents a human LGR6 polypeptide.

Sequence 633 AA:

Query Match 99.9%; Score 3305; DB 23; Length 633;
 Best Local Similarity 100.0%; Pred. No. 2,66-293;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTHRESWYACRYSRSGTSHASVRSQGLSLPAHPASIALAASNTTASGKLEXTF 60
 DB 1 NTHRESWYACRYSRSGTSHASVRSQGLSLPAHPASIALAASNTTASGKLEXTF 60
 QY 61 SOLSLQALDLSWNAIRSIHPEAFSTLSLVKLDLTNDQTLPLAGLGLMHLKGNL 120
 DB 61 SOLSLQALDLSWNAIRSIHPEAFSTLSLVKLDLTNDQTLPLAGLGLMHLKGNL 120
 QY 121 ALSQAFKSDSPKRILEVPAVYOCCEPGKASFGQWEADLHDDDESSKRPGL 180
 DB 121 ALSQAFKSDSPKRILEVPAVYOCCEPGKASFGQWEADLHDDDESSKRPGL 180
 QY 181 LARQENHYDDDELQLEMEDSKRHPVOCSPNPPKPEYLFESKGRFLAWAIVL 240
 DB 181 LARQENHYDDDELQLEMEDSKRHPVOCSPNPPKPEYLFESKGRFLAWAIVL 240
 QY 241 SVLCNGVILTVRAGGAPLPVYKVVGAAGANTLGTISGGLASVADALTFQGFSEGA 300
 DB 241 SVLCNGVILTVRAGGAPLPVYKVVGAAGANTLGTISGGLASVADALTFQGFSEGA 300
 QY 301 RWETGLGCRATGFLAVLGSEASVLLTLTAVVOCSSVSCVRAGKSPSLCSVRAGVLC 360
 DB 301 RWETGLGCRATGFLAVLGSEASVLLTLTAVVOCSSVSCVRAGKSPSLCSVRAGVLC 360
 QY 361 ATAGLAAALPLASVGEYASPLCPYAPPEGGPAAAGFTVALVMNMSFCLVAVAGATK 420
 DB 361 ATAGLAAALPLASVGEYASPLCPYAPPEGGPAAAGFTVALVMNMSFCLVAVAGATK 420
 QY 421 YCDLPGRGDEAVWDCAMVHVAMLIIFADGLICPVAFLSPASMLGLFVTPPEAVKSVLLV 480
 DB 421 YCDLPGRGDEAVWDCAMVHVAMLIIFADGLICPVAFLSPASMLGLFVTPPEAVKSVLLV 480

QY 481 VLPLPACLNPLLYLLFNPNPRDDLRRLRPAGDSGLAAYAAAGELKSSCDSTQALVARS 540
 DB 481 VLPLPACLNPLLYLLFNPNPRDDLRRLRPAGDSGLAAYAAAGELKSSCDSTQALVARS 540
 QY 541 DVDLLLEASAGRPGLERYGFPVSVTLISCOQPGAPRLGSGHCVEEGNHFQPSMDG 600
 DB 541 DVDLLLEASAGRPGLERYGFPVSVTLISCOQPGAPRLGSGHCVEEGNHFQPSMDG 600
 QY 601 ELLLRAGSTPAGGGLSGGGGFPQPSGLAFASHV 633
 DB 601 ELLLRAGSTPAGGGLSGGGGFPQPSGLAFASHV 633

RESULT 2
 AAG66140
 ID AAG66140 standard; Protein: 736 AA.
 XX
 AC AAG66140;
 XX
 DT 13-MAR-2002 (first entry)
 XX
 DE Human LGR6 polypeptide (clone fahr).

XX Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
 KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
 KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiac;
 XX antiarrhythmic; anorectic; gene therapy; human.

OS Homo sapiens.

PN WO200185768-A2.

PD 15-NOV-2001.

PF 08-MAY-2001; 2001WO-US15002.

PR 08-MAY-2000; 2000US-0566588.

PA (MILL-) MILLENNIUM PHARM INC.

PI Gu W;

DR WPI: 2002-055584/07.

DR N-PSDB: AA167924, AA167925.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 PT disorder

Claim 9; Fig 8; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular process, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypothyroidism, hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of

CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a human LG6 polypeptide.

SO Sequence 736 AA;

Query Match 91.2%; Score 3015; DB 23; Length 736;

Best Local Similarity 99.3%; Pred. No. 1.1e-266; Mismatches 2; Indels 0; Gaps 0;
 Matches 576; Conservative

QY 54 KLEADTFSSQLADLSWNAIRSHPEAFSTHSLVKLDITDNOITPLAGLGIMH 113
 DB 157 EIGADTFSSQLADLSWNAIRSHPEAFSTHSLVKLDITDNOITPLAGLGIMH 216
 QY 114 LKLGKGLALSAFESKSPFKRLILEVYAYOCPCYGMKASFFKASQWEADLHLDDEES 173
 DB 217 LKLGKGLALSAFESKSPFKRLILEVYAYOCPCYGMKASFFKASQWEADLHLDDEES 276
 QY 174 SKRPLGLARQAEHNYDDDELQLEMEDESKRPHSVQCSPTPGPFCEYLFESWGIRLA 233
 DB 277 SKRPLGLARQAEHNYDDDELQLEMEDESKRPHSVQCSPTPGPFCEYLFESWGIRLA 336
 QY 234 VMAIVLLSVLCNGVLITVFAAGPAPLPVKFVGAIGANTLTGISCGLASVDALTFG 293
 DB 337 VMAIVLLSVLCNGVLITVFAAGPAPLPVKFVGAIGANTLTGISCGLASVDALTFG 396
 QY 294 QFSEYGARWETGLGCRATGFLAVLGSEASVLLTLAVQCSVSYSCYRAVKSPLGSVR 353
 DB 397 QFSEYGARWETGLGCRATGFLAVLGSEASVLLTLAVQCSVSYSCYRAVKSPLGSVR 456
 QY 354 AGVIGCLALAGLAAALPLASVGEYASPLCPYAPPEGQPALGFTVALVMNSFCFLV 413
 DB 457 AGVIGCLALAGLAAALPLASVGEYASPLCPYAPPEGQPALGFTVALVMNSFCFLV 516
 QY 414 AGAIVKLYCDLPRGDFEAVMDCAMVRHVAWLIIFADGLLYCPVAIFSPSMGIFPTEPA 473
 DB 517 AGAIVKLYCDLPRGDFEAVMDCAMVRHVAWLIIFADGLLYCPVAIFSPSMGIFPTEPA 576
 QY 474 VKSVLLVPLPACINPLLYLTFNPFRRDLRLRRAGDSGPLAYAAAGELKSSCDST 533
 DB 577 VKSVLLVPLPACINPLLYLTFNPFRRDLRLRRAGDSGPLAYAAAGELKSSCDST 636
 QY 534 QALVAFSDVDLLEASSEAGRPPELTGFPSPVTLISCCQPGARLEGSCVPEGNHGN 593
 DB 637 QALVAFSDVDLLEASSEAGRPPELTGFPSPVTLISCCQPGARLEGSCVPEGNHGN 696
 QY 594 POPSMDELILRAEGSTPAGGSLSGGFGPSPGIAFASHV 633
 DB 697 POPSMDELILRAEGSTPAGGSLSGGFGPSPGIAFASHV 736

RESULT 3
 AAE23418
 ID AAE23418 standard; Protein: 948 AA.
 AC AAE23418;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 XX Human G-protein coupled receptor-10 (GCRG-10) protein.
 XX
 XX Human: G-protein coupled receptor-10; GCRG-10; atherosclerosis; cancer;
 KM cell proliferative disorder; gastrointestinal; autoimmune; metabolic;
 KM neurological; inflammatory; cardiovascular; viral infection; anorexia;
 KM cirrhosis; multiple sclerosis; Huntington's disease; gene therapy;
 KM Alzheimer's disease; heart failure; hepatitis; cholecystitis; obesity;
 KM rheumatoid arthritis; Hashimoto's thyroiditis; angina pectoris;
 KM osteoporosis.
 KM
 XX Homo sapiens.
 OS
 XX
 XX
 XX Key 1.25 location/Qualifiers
 FT Peptide /label= signal_peptide

FT Protein 26..948
 FT /note= "Mature GCRG-10 protein"
 FT 34..65 "Leucine rich repeat N-terminal domain"
 FT /note= 548..571
 FT Domain /label= transmembrane_domain
 FT 756..783
 FT Domain /label= transmembrane_domain
 PN W0200226825-A2.
 PD 04-APR-2002.
 XX
 XX
 PF 28-SEP-2001; 2001WO-US30661.
 XX
 PR 29-SEP-2000; 2000US-236546P.
 PR 13-OCT-2000; 2000US-240589P.
 PR 20-OCT-2000; 2000US-242232P.
 PR 20-OCT-2000; 2000US-242322P.
 PR 03-NOV-2000; 2000US-245835P.
 PR 03-NOV-2000; 2000US-245900P.
 PR 09-NOV-2000; 2000US-247587P.
 PR 15-NOV-2000; 2000US-249343P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Baughn MR, Graul RC, Walla NK, Gandhi AR, Hatfield AJ, Rankumar J;
 PI Tribouley CM, Thornton M, Kallick DA, Yao MG, Elliott VS;
 PI Burford N, Khan FA, Yue H, Lu Y, Arvizu C, Roopa R, Nguyen DB;
 PI Lee EA, Lu DAM, Ison CH, Walsh RT, Policky JT;
 XX
 XX WPI: 2002-426012/45.
 DR N-PSDB: AAD37674.
 XX
 XX Novel G-protein coupled receptor polypeptides referred as GCRG
 PI peptides, useful for treating multiple sclerosis, cholecystitis heart
 PT failure, angina pectoris, rheumatoid arthritis, obesity, osteoporosis
 PT
 PT
 PT
 PS Claim 71; Page 126-129; 147pp; English.
 XX
 XX The invention relates to human G-protein coupled receptor (GCRG 1-16)
 CC polypeptides and polynucleotides. Sequences of the invention are useful
 CC for the diagnosis, treatment and prevention of cell proliferative (e.g.
 CC cancer, cirrhosis), neurological (e.g. multiple sclerosis, Huntington's
 CC disease, Alzheimer's disease), cardiovascular (e.g. angina pectoris,
 CC heart failure), gastrointestinal (e.g. anorexia, cholecystitis),
 CC autoimmune/inflammatory (rheumatoid arthritis, Hashimoto's thyroiditis)
 CC and metabolic disorders (e.g. obesity, osteoporosis), viral infections,
 CC atherosclerosis and hepatitis. GCRG proteins are useful for identifying
 CC compounds that modulate, mimic and block olfactory and taste sensations.
 CC They are also useful for identifying GCRG modulators. GCRG DNAs are
 CC useful in gene therapy. The present sequence is human GCRG-10 protein.
 CC
 CC
 XX
 XX
 SO Sequence 948 AA;
 Query Match 91.2%; Score 3015; DB 23; Length 948;
 Best Local Similarity 99.3%; Pred. No. 1.6e-266; Mismatches 2; Indels 0; Gaps 0;
 Matches 576; Conservative

QY 54 KLEADTFSSQLADLSWNAIRSHPEAFSTHSLVKLDITDNOITPLAGLGIMH 113
 DB 369 EIGADTFSSQLADLSWNAIRSHPEAFSTHSLVKLDITDNOITPLAGLGIMH 428
 QY 114 LKLGKGLALSAFESKSPFKRLILEVYAYOCPCYGMKASFFKASQWEADLHLDDEES 173
 DB 429 LKLGKGLALSAFESKSPFKRLILEVYAYOCPCYGMKASFFKASQWEADLHLDDEES 488
 QY 174 SKRPLGLARQAEHNYDDDELQLEMEDESKRPHSVQCSPTPGPFCEYLFESWGIRLA 233
 DB 489 SKRPLGLARQAEHNYDDDELQLEMEDESKRPHSVQCSPTPGPFCEYLFESWGIRLA 548
 QY 234 VMAIVLLSVLCNGVLITVFAAGPAPLPVKFVGAIGANTLTGISCGLASVDALTFG 293

```

Db 549 VVAIVLVLSTVLCNGVLLVTFVAGGPAFLPVKTVGAIAGANTLTGISCGLASVDALITFG 608
QY 294 QPSEYGARWETGLGCRATGFLAVLGSSEASVLLTFLAAVOCSSVSCVRAYGKSPSLGSVR 353
Db 609 QPSEYGARWETGLGCRATGFLAVLGSSEASVLLTFLAAVOCSSVSCVRAYGKSPSLGSVR 668
QY 354 AGVLGCLALAGLAALPLASVGEYASPLCLPYAPBEGPALGFTVALVMNNSFCFLVY 413
Db 669 AGVLGCLALAGLAALPLASVGEYASPLCLPYAPBEGPALGFTVALVMNNSFCFLVY 728
QY 414 AGATIKLYCDLPBGDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 473
Db 729 AGATIKLYCDLPBGDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 788
QY 474 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 533
Db 789 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 848
QY 534 QALVAFSDVDLILEASERGPGLGTYGFPSSVTLISCOQPGAPRLGSHCVPEPGNHNFGN 593
Db 849 QALVAFSDVDLILEASERGPGLGTYGFPSSVTLISCOQPGAPRLGSHCVPEPGNHNFGN 908
QY 594 POPSMDELLLRAGSTPAGGSLGGGGFOPSGIAFAASHV 633
Db 909 POPSMDELLLRAGSTPAGGSLGGGGFOPSGIAFAASHV 948

RESULT 4
AAG66141
ID AAG66141 standard; Protein; 967 AA.
XX
AC AAG66141;
XX
DT 13-MAR-2002 (first entry)
XX
DE Human LGR6 polypeptide (clone Fb150881).
XX
KM Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
KM anti-HIV; cytostatic; neurotropic; neuroprotective; antiparkinsonian;
KM anticonvulsant; neuroleptic; antidepressant; tranquilizer; antidiabetic;
KM osteopontin; antihypertensive; vasodilator; hypotensive; cardiac;
KM antitachycardic; anorectic; gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO200185768-A2.
XX
PD 15-NOV-2001.
XX
PE 08-MAY-2001; 2001WO-US15002.
XX
PR 08-MAY-2000; 2000US-0566588.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
DR WPI: 2002-055584/07.
XX
N-PSDB; AAI67926, AAI67927.
XX
PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
PT for diagnosing and treating weight disorder, metabolic disorder,
PT central nervous system disorder, endocrine disorder and cardiovascular
PT disorder.
XX
PS Claim 9; Fig 15; 198pp; English.
XX
CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC polynucleotides are useful as targets for developing modulating agents and
CC processes, as well as thermogenesis. They are useful for developing novel

```

```

CC diagnostic and therapeutic agents for LGR6 associated disorders such as
CC weight disorders (anorexia, obesity), eating disorders, acquired
CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, Gilles de la
CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
CC endocrine disorders (hypothalamic and hypothalamic disorders, dwarfism, gigantism,
CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC polynucleotides and antibodies are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g., diagnostic assays, prognostic assays,
CC monitoring clinical trials and pharmacogenomics), and in methods of
CC treatment (e.g., therapeutic and prophylactic). The present sequence
XX represents a human LGR6 polypeptide.
SQ Sequence 967 AA;

```

```

Query Match 91.2%; Score 3015; DB 23; Length 967;
Best Local Similarity 99.3%; Pred. No. 1.6e-266;
Matches 576; Conservative 2; Mismatches 2; Indels 0; Caps 0;

```

```

QY 54 KLENDTFSQSLSSQALDLSMNAIRSIHPEAFSTLSLVKLDITDNOQTTPPLAGLGIMH 113
Db 388 ELGADTFSSQLSSQALDLSMNAIRSIHPEAFSTLSLVKLDITDNOQTTPPLAGLGIMH 447
QY 114 LKLGKLNALSQAFSKDSEFPKRLILEVYAYOCGCGYGCACAFRRASGWEADJLHDEES 173
Db 448 LKLGKLNALSQAFSKDSEFPKRLILEVYAYOCGCGYGCACAFRRASGWEADJLHDEES 507
QY 174 SKRPLGLLAQENHNDQIDDELQLEMEDSKPHPSVQCSPTPEPFKCELYFESWGRILA 233
Db 508 SKRPLGLLAQENHNDQIDDELQLEMEDSKPHPSVQCSPTPEPFKCELYFESWGRILA 567
QY 234 VVAIVLVLSTVLCNGVLLVTFVAGGPAFLPVKTVGAIAGANTLTGISCGLASVDALITFG 293
Db 568 VVAIVLVLSTVLCNGVLLVTFVAGGPAFLPVKTVGAIAGANTLTGISCGLASVDALITFG 627
QY 294 QPSEYGARWETGLGCRATGFLAVLGSSEASVLLTFLAAVOCSSVSCVRAYGKSPSLGSVR 353
Db 628 QPSEYGARWETGLGCRATGFLAVLGSSEASVLLTFLAAVOCSSVSCVRAYGKSPSLGSVR 687
QY 354 AGVLGCLALAGLAALPLASVGEYASPLCLPYAPBEGPALGFTVALVMNNSFCFLVY 413
Db 688 AGVLGCLALAGLAALPLASVGEYASPLCLPYAPBEGPALGFTVALVMNNSFCFLVY 747
QY 414 AGATIKLYCDLPBGDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 473
Db 748 AGATIKLYCDLPBGDFEAVWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGLFPVTEPA 807
QY 474 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 533
Db 808 VKSVLLVVLPLPACLNPLLYLTFNPHFRDDLRRLRPRAAGSGPLAAYAAAGLEKSSCDST 867
QY 534 QALVAFSDVDLILEASERGPGLGTYGFPSSVTLISCOQPGAPRLGSHCVPEPGNHNFGN 593
Db 868 QALVAFSDVDLILEASERGPGLGTYGFPSSVTLISCOQPGAPRLGSHCVPEPGNHNFGN 927
QY 594 POPSMDELLLRAGSTPAGGSLGGGGFOPSGIAFAASHV 633
Db 928 POPSMDELLLRAGSTPAGGSLGGGGFOPSGIAFAASHV 967

RESULT 5
AAV53575
ID AAV53575 standard; Protein; 847 AA.
XX
AC AAV53575;
XX
DT 15-FEB-2000 (first entry)
XX

```

D6	609 AGATIKLYCDLPGRGDEFAVMDCAVRHYVALMLPADGLLTYCPVAFLFSAASLMGLFPVTPEA	668
OY	474 VKSVLLVYLPLPALINLILYLENPHFRDDLRLRPARGSGPLATAAAGLEKSSCDST	533
D6	669 VKSVLVYLPLPALINLILYLENPHFRDDLRLRPARGSGPLATAAAGLEKSSCDST	728
OY	534 QALVFSDVDILLEASEGRPGLTTFYSPSVTLISCOQPGAPRIECSHCVEPCGNHFGN	593
D6	729 QALVAFSDVDILLEASEGRPGLTTFYSPSVTLISCOQPGAPRIECSHCVEPCGNHFGN	788
OY	594 POPSMDELLLRAREGSTPRAGGSLGGGPFOP	624
D6	789 POPSMDELLLRAREGSTPRAGGSLGGGAFSP	819
<hr/>		
RESULT 6		
ID	AAU79941 standard; Protein; 940 AA.	
XX	AAU79941:	
AC		
XX	15-JUL-2002 (first entry)	
DT		
XX	Human orphan GPCR-like protein NOV3.	
DE		
XX	Human: NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;	
KW	cell signal processing; metabolic pathway modulation; cancerous tissue;	
KM	antibody; diabetes; transgenic animal; orphan GPCR-like protein; NOV3.	
KN		
XX	Homo sapiens.	
OS		
XX	WO200229038-A2.	
PN		
XX	11-APR-2002.	
PD		
XX	04-OCT-2001; 2001WO-US31377.	
PF		
XX	04-OCT-2000; 2000US-237862P.	
PR		
XX	(CURA-) CURAGEN CORP.	
PA		
XX	Herrmann JL, Rastelli L, Shinkets RA;	
PI		
XX	WP1; 2002-340104/37.	
DR	N-PSDB; ABK49424.	
XX		
PT	Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for	
PT	treating cardiomyopathy, atherosclerosis, and cancer -	
XX	Claim 1; Page 36; 180pp; English.	
PS		
XX	The present invention relates to a new NOVX polypeptide having a	
CC	900 (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or	
CC	331 (NOV6) residue amino acid sequence, as given in the specification.	
CC	The novel polypeptide, and its encoding polynucleotide, are used to	
CC	treat cardiomyopathy, atherosclerosis, cancer or a disease related to	
CC	cell signal processing and metabolic pathway modulation, in a human	
CC	cancerous tissue. The antibody can be used to treat diabetes or cancer.	
CC	The host cells can be used to produce non-human transgenic animals	
CC	useful in drug screening. The present amino acid sequence represents	
CC	the human orphan GPCR-like protein NOV3 of the invention.	
XX		
SQ	Sequence 940 AA:	
<hr/>		
Query Match	89.5%; Score 2960; DB 23; Length 940;	
Best Local Similarity	98.9%; Pred. No. 1.7e+261;	
Matches 565; Conservative	2; Mismatches 4; Indels 0; Gaps 0;	
OY	54 KLEADTSOQLSQAOLDLSMNATSIHPFASTLHSIVKLDITDNLQTTPLAGIGIMH	113
D6	342 EIGADTFQSLSQALDDLSMNAINSHIPEASTLHSIVKLDITDNLQTTPLAGIGIMH	401

QY 114 LKLGNTALSOAFKSDSEPKRLILEVPAAYOCCPYGMKASFFKASGQWAEADLHLDDEES 173
 DB 402 LKLGNTALSOAFKSDSEPKRLILEVPAAYOCCPYGMKASFFKASGQWAEADLHLDDEES 461
 QY 174 SKRPLGLARQAEHNYDQDDELQLEMEDESKRPSVOCSPPGPFKCEYLFESWGIRLA 233
 DB 462 SKRPLGLARQAEHNYDQDDELQLEMEDESKRPSVOCSPPGPFKCEYLFESWGIRLA 521
 QY 234 VMAIVLLSVLCNGILVLTVPAGPAPLPYKPVVGAIAAGANTLTGISCGLASVDALTFEG 293
 DB 522 VMAIVLLSVLCNGILVLTVPAGPAPLPYKPVVGAIAAGANTLTGISCGLASVDALTFEG 581
 QY 294 QFSEYGARWETGLGCRATGTLAVLSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
 DB 582 QFSEYGARWETGLGCRATGTLAVLSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 641
 QY 354 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 413
 DB 642 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 701
 QY 414 AGAYIKLYCDLPRGDEFAVNDCAVNRHVAMLIADGLLYCPVAFSLFASMLGLFVTPPEA 473
 DB 702 AGAYIKLYCDLPRGDEFAVNDCAVNRHVAMLIADGLLYCPVAFSLFASMLGLFVTPPEA 761
 QY 474 VKSULLVLPPLACLNPLLYLTFNPHFRDRLRLRPRAGDSGLAYAAAGLEKSSCDST 533
 DB 762 VKSULLVLPPLACLNPLLYLTFNPHFRDRLRLRPRAGDSGLAYAAAGLEKSSCDST 821
 QY 534 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPEGNHFGN 593
 DB 822 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPEGNHFGN 881
 QY 594 POPMDELLLRAGSTIPACGGSLSGGGGFP 624
 DB 882 POPMDELLLRAGSTIPACGGSLSGGGGFP 912

RESULT 7

AA66138
 ID AAG66138 standard; Protein: 967 AA.
 AC AAG66138;

DT 13-MAR-2002 (first entry)
 DE Mouse LGR6 polypeptide (clone flmzb048h10).

KW Large G-protein coupled receptor 6; LGR6; GPCR; Immunomodulator;
 KW anti-HIV; cytoskeletal; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antiemigraine;
 KW osteopathic; antiarteriosclerotic; vasodilator; hypotensive; cardiant;
 KW antiarhythmic; anorectic; gene therapy; mouse.
 OS Mus musculus.

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note= "signal peptide"
 FT /note= "24..967
 FT /note= "mature protein"

PN MO200185768-A2.
 PD 15-NOV-2001.

PF 08-MAY-2001; 2001WO-US15002.
 PR 08-MAY-2000; 2000US-0566588.

PA (MILL-) MILLENNIUM PHARM INC.
 PL Gu W;
 XX

DR WPI: 2002-055584/07.
 DR N-PSDB: AAI67920, AAI67921.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 PT disorder
 PS Example 1; Fig 1; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC obsessive compulsive disorder, anxiety, panic disorder, schizophrenia,
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g., diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g., therapeutic and prophylactic). The present sequence
 CC represents a mouse LGR6 polypeptide.
 SQ Sequence 967 AA;

Query Match 79.7%; Score 2637; DB 23; Length 967;
 Best Local Similarity 87.2%; Pred. No. 6.3e-232;
 Matches 506; Conservative 28; Mismatches 46; Indels 0; Gaps 0;

QY 54 LKENDTSQSLQADLISWNAIRSIHPEAFSTLSLVKLDLNDQTLTPPLAGLGLMH 113
 DB 388 EIGADYFSQSLQADLISWNAIRSIHPEAFSTLSLVKLDLNDQTLTPPLAGLGLMH 113
 QY 114 LKLGNTALSOAFKSDSEPKRLILEVPAAYOCCPYGMKASFFKASGQWAEADLHLDDEES 173
 DB 448 LKLGNTALSOAFKSDSEPKRLILEVPAAYOCCPYGMKASFFKASGQWAEADLHLDDEES 173
 QY 174 SKRPLGLARQAEHNYDQDDELQLEMEDESKRPSVOCSPPGPFKCEYLFESWGIRLA 233
 DB 508 SKRPLGLARQAEHNYDQDDELQLEMEDESKRPSVOCSPPGPFKCEYLFESWGIRLA 233
 QY 234 VMAIVLLSVLCNGILVLTVPAGPAPLPYKPVVGAIAAGANTLTGISCGLASVDALTFEG 293
 DB 568 VMAIVLLSVLCNGILVLTVPAGPAPLPYKPVVGAIAAGANTLTGISCGLASVDALTFEG 293
 QY 294 QFSEYGARWETGLGCRATGTLAVLSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
 DB 628 QFSEYGARWETGLGCRATGTLAVLSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSVR 353
 QY 354 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 413
 DB 688 AGVLGCLALAGLAAALPLASVGEYASPLCLPYAPPEGQPALGFTVALVMNNSFCFLV 413
 QY 414 AGAYIKLYCDLPRGDEFAVNDCAVNRHVAMLIADGLLYCPVAFSLFASMLGLFVTPPEA 473
 DB 748 AGAYIKLYCDLPRGDEFAVNDCAVNRHVAMLIADGLLYCPVAFSLFASMLGLFVTPPEA 473
 QY 474 VKSULLVLPPLACLNPLLYLTFNPHFRDRLRLRPRAGDSGLAYAAAGLEKSSCDST 533
 DB 808 VKSULLVLPPLACLNPLLYLTFNPHFRDRLRLRPRAGDSGLAYAAAGLEKSSCDST 533
 QY 534 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPEGNHFGN 593
 DB 867 QALVAFSDVDLILEASEAGRPGLTETGFPSTVLTISCCQPGAPRLKESHCVEPEGNHFGN 593

Query Match	75.4%	Score 2494.5	DB 22	Length 691
Best Local Similarity	87.0%	Score No. 4.3e-219		
Matches 504; Conservative	12	Mismatches 50	Indels 13	Gaps 9

AA
DB WPT: 1999-563673/48

XX Van Der Spek PJ, Heikoop JC;
PI
XX WPT, 1999-563673/48.
DR N-PSDB; AA240460.
DR

XX New 7 transmembrane gonadotropin receptors, useful for screening for
PT hormone analogs and drugs -
PS Claim 1, Page 18-21, 38pp; English.

CC Sequences AAY5371-Y53578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.

Sequence 497 AA:

Query Match 74.1%; Score 2450; DB 20; Length 497;
Best Local Similarity 99.6%; Pred. No. 3.2e-215;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

156 KASGQAEADLHDDDESSKRLGLARQAEHNDLDELQLEMEDSKPSPVQCSPTP 215
- 1 KASGQAEADLHDDDESSKRLGLARQAEHNDLDELQLEMEDSKPSPVQCSPTP 60
216 GPPKPEEYLFESMGIRLAWAIVLVSVLNGVLLTVFAGGAPAPLPVKFVGAAGANT 275
61 GPPKPEEYLFESMGIRLAWAIVLVSVLNGVLLTVFAGGAPAPLPVKFVGAAGANT 120
276 LKSGISGLASVDALTFGGFSESGARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSV 335
121 LKSGISGLASVDALTFGGFSESGARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSV 180
336 SVSCVRANGKSPSLGSRVAGVGLCLALAGLAALPLASVEGASPLCLPYAPPEGPA 395
181 SVSCVRANGKSPSLGSRVAGVGLCLALAGLAALPLASVEGASPLCLPYAPPEGPA 240
396 LGFTVALVWMSFCLVAVAGATIKLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPV 455
241 LGFTVALVWMSFCLVAVAGATIKLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPV 300
456 AFLSFASMLGLFPVTPPEAVKSVLLVPLPACINPLLYLLENFHRDDLRLRPRAGDSG 515
301 AFLSFASMLGLFPVTPPEAVKSVLLVPLPACINPLLYLLENFHRDDLRLRPRAGDSG 360
516 PLAAAGAELEKSSCDSTQALVAFSDVDLLEASERGRPGLETYGFPSVTLISCOOPGA 575
361 PLAAAGAELEKSSCDSTQALVAFSDVDLLEASERGRPGLETYGFPSVTLISCOOPGA 420
576 PRLGSHQVEPEGNHFGNPPSMOGEILLRAGSTPAGGGLSGGGGFP 624
421 PRLGSHQVEPEGNHFGNPPSMOGEILLRAGSTPAGGGLSGGGGFP 469

RESULT 10

ID AAG67556 standard; Protein; 395 AA.

XX AAG67556;

DT 26-NOV-2001 (first entry)

XX A human guanosine triphosphate binding protein coupled receptor.

KM Human; guanosine triphosphate binding protein coupled receptor;

OS G protein coupled receptor; brain disease; cancer.

XX Homo sapiens.

XX WO200109323-A1.

PD 08-FEB-2001.

XX 28-JUL-2000; 2000WO-JP05070.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 18-OCT-1999; 99JP-0159590.

XX 11-JAN-2000; 2000JP-0118776.

XX 17-FEB-2000; 2000US-0183322.

XX 02-MAY-2000; 2000JP-0183767.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Kishimoto T;

XX Yano K, Kanazaki K, Inoue Y;

XX WPI; 2001-570288/64.

XX N-PSDB; AAH78275.

XX New gene encoding guanosine triphosphate binding protein coupled

XX receptor, and the protein and antibodies to it, useful for diagnosis

XX and treatment of disease such as brain disease -

XX Claim 2; Page 48-51; 63pp; Japanese.

XX The present sequence represents a human guanosine triphosphate binding

XX protein coupled receptor. The guanosine triphosphate binding

XX protein coupled receptor protein is useful in the diagnosis, prediction

XX and treatment of disease associated with disorders of G protein coupled

XX receptor protein, and may be useful in brain disease and cancers.

XX Sequence 395 AA:

Query Match 61.5%; Score 2033; DB 22; Length 395;
Best Local Similarity 99.5%; Pred. No. 3.3e-177;
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

239 LLSVLCNGLVLLTVFAGGAPAPLPVKFVGAAGANTLGLISGLASVDALTFGGFSEY 298
1 LLSVLCNGLVLLTVFAGGAPAPLPVKFVGAAGANTLGLISGLASVDALTFGGFSEY 60

299 GARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSVSCVRANGKSPSLGSRVAGVIG 358
61 GARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSVSCVRANGKSPSLGSRVAGVIG 120

359 CLALAGLAALPLASVEGASPLCLPYAPPEGOPALGFTVALVWMSFCLVAVAGAT 418
121 CLALAGLAALPLASVEGASPLCLPYAPPEGOPALGFTVALVWMSFCLVAVAGAT 180

419 KLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPVAFLSFASMLGLFPVTPPEAVKSVL 478
181 KLYCDLPRGDFEAVWDCAMVRHVAWLTFADGLLYCPVAFLSFASMLGLFPVTPPEAVKSVL 240

479 LVVLPPLACINPLLYLLENFHRDDLRLRPRAGDSGLPAAAGAELEKSSCDSTQALVA 538
241 LVVLPPLACINPLLYLLENFHRDDLRLRPRAGDSGLPAAAGAELEKSSCDSTQALVA 300

539 FSDVDLLEASERGRPGLETYGFPSVTLISCOOPGAPRLGSHQVEPEGNHFGNPPSM 598
301 FSDVDLLEASERGRPGLETYGFPSVTLISCOOPGAPRLGSHQVEPEGNHFGNPPSM 360

599 DGEILLRABGSTPAGGGLSGGGGFPSPGLAFASHV 633
361 DGEILLRABGSTPAGGGLSGGGGFPSPGLAFASHV 395

RESULT 11

ID AAW93889 standard; Protein; 907 AA.

XX AAW93889

XX AAW93889;

[illegible]

Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;

```

OY 54 KLEXTFESQSLQALDLSWNAISIRHPEAFSTLSVTKLDLTDNQTLTTLPLAGLGIMH 113
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 388 EIKVDTEQQLLSIRSLINAMKIAIIRHNAFSTLSPLTKLDLSNLSLSPFITGHLGTH 447
OY 114 LKLGKLNALSGAFSKDSFPKRLILEVPYAYOCPCPYGMCAFPKASGOW-----BAEDLH 167
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 448 LKLGKLNALSGAFSKDSFPKRLILEVPYAYOCPCPYGMCAFPKASGOW-----BAEDLH 167
OY 168 LDDESSKRPGLGLARQAEHNYDDLDLEDEM-EDSKPHPSVOCSPTPGPFKCEYLFE 226
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 508 -----KRDAGMFOADE-----RDLEDFLLDEEDLKAHLSVOCSPSPGPFKCEHLLD 556
OY 227 SWGIRLAWAIVLVSLVNCGLVLTVEFAGGAPALPYKFEVVGALAGNTLTGISCGLAS 286
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 557 GMLIRIGWMTIAVLTALCNALVTSTVER-SPLYISPIKLIGVIAAVNMLTGVSAAVLG 615
OY 287 VDALTFGQFSEYGARWETGLGRATGFLAVLGSEASVLLTLTAAVOCSSVSCVRAVYGS 346
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 616 VDAFTFGSFARHGAMWENGCHVIGFLSIFASESSVFLLTLAALRGFSVKYSAKFEYK 675
OY 347 PSLGVRAGVIGCLALAGLAALPLASVEGYGASPLCLPYAPPEGAPALGFTVALVMN 406
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 676 APFSSILKVIILLCALLALTMMAVPLLGSKYASPLCLPL--PFGEPSTMGVWVALITLN 733
OY 407 SFCEFLVAGAVYIKLYCDLPGRGFEAVWDCAMVRYHVAMLIFADGLTLCVPAFLSFAAMLGL 466
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 734 SLCEFLMTIATYTKLCNLDKGDLENIMDCSMVKHIALLLFTNCILNCPVAFLESSESLNL 793
OY 467 FPVTPRAVKSVLVPLPACINPLLYLLENPHFRDRLRLRPRA-----GDSGPLAYAA 521
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 794 TFIISPEVIRKFIILVVPPLPACINPLLYLLENPHFRDRLRLRPRA-----GDSGPLAYAA 521
OY 522 AGELEKSSCDSTQALVAFSDVDLLEASEAGRP-----PGLFTYGPSPSYTLISC 570
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 854 SDOVEKQSCDSTQALVFTSSSITVDLPSSSVPSPAYVTESCHLSVAFAVPC 906

```

RESULT 13

ID AAY42169 standard; Protein: 693 AA.

```

XX AAY42169;
AC
XX 20-DEC-1999 (first entry)
DT
XX Human LGR5 protein sequence.
DE
XX Human LGR5 protein sequence.
KW Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
XX extracellular leucine rich repeat region; mapping; identification.
OS Homo sapiens.
XX
XX WO9948921-A1.
XX
XX 30-SEP-1999.
PD
XX 25-MAR-1999; 99WO-US06573.
PF
XX 26-MAR-1998; 98US-0079501.
PR
XX
XX (STRD ) UNIV LELAND STANFORD JUNIOR.
PA
XX (ORCA ) ORGANON NV.
PI
XX Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX WPI; 1999-591074/50.
XX N-PSDB; AAZ5344.
XX
XX New G-protein coupled receptors, useful for identifying their own
XX ligands -
XX

```

PS Claim 2; Fig 2; 54pp; English.

XX The present sequence represents the human G-protein coupled receptor
 CC having extracellular leucine rich repeat regions, designated LGR5.
 CC The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
 CC the receptor. The polypeptides and/or polynucleotides are also useful
 CC for homologous or related genes, producing compositions that modulate
 CC the expression or function of the receptors, gene therapy, mapping
 CC functional regions of the receptors, studying associated physiological
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
 CC for producing antibodies, and for identifying biologically active
 CC agents. The polypeptides contain a G-protein coupled seven
 CC transmembrane region and a leucine rich repeat extracellular domain.
 CC These regions capture and facilitate optimal orientation of its ligand.
 CC The proteins are also expressed in diverse tissues.

SO Sequence 693 AA;

Query Match 38.5%; Score 1273; DB 20; Length 693;
 Best Local Similarity 46.9%; Pred. No. 2.6e-107;
 Matches 250; Conservative 96; Mismatches 157; Indels 30; Gaps 8;

```

OY 54 KLEXTFESQSLQALDLSWNAISIRHPEAFSTLSVTKLDLTDNQTLTTLPLAGLGIMH 113
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 174 EIKVDTEQQLLSIRSLINAMKIAIIRHNAFSTLSPLTKLDLSNLSLSPFITGHLGTH 233
OY 114 LKLGKLNALSGAFSKDSFPKRLILEVPYAYOCPCPYGMCAFPKASGOW-----BAEDLH 167
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 234 LKLGKLNALSGAFSKDSFPKRLILEVPYAYOCPCPYGMCAFPKASGOW-----BAEDLH 167
OY 168 LDDESSKRPGLGLARQAEHNYDDLDLEDEM-EDSKPHPSVOCSPTPGPFKCEYLFE 226
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 294 -----KRDAGMFOADE-----RDLEDFLLDEEDLKAHLSVOCSPSPGPFKCEHLLD 342
OY 227 SWGIRLAWAIVLVSLVNCGLVLTVEFAGGAPALPYKFEVVGALAGNTLTGISCGLAS 286
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 343 GMLIRIGWMTIAVLTALCNALVTSTVER-SPLYISPIKLIGVIAAVNMLTGVSAAVLG 401
OY 287 VDALTFGQFSEYGARWETGLGRATGFLAVLGSEASVLLTLTAAVOCSSVSCVRAVYGS 346
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 402 VDAFTFGSFARHGAMWENGCHVIGFLSIFASESSVFLLTLAALRGFSVKYSAKFEYK 461
OY 347 PSLGVRAGVIGCLALAGLAALPLASVEGYGASPLCLPYAPPEGAPALGFTVALVMN 406
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 462 APFSSILKVIILLCALLALTMMAVPLLGSKYASPLCLPL--PFGEPSTMGVWVALITLN 519
OY 407 SFCEFLVAGAVYIKLYCDLPGRGFEAVWDCAMVRYHVAMLIFADGLTLCVPAFLSFAAMLGL 466
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 520 SLCEFLMTIATYTKLCNLDKGDLENIMDCSMVKHIALLLFTNCILNCPVAFLESSESLNL 579
OY 467 FPVTPRAVKSVLVPLPACINPLLYLLENPHFRDRLRLRPRA-----GDSGPLAYAA 521
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 580 TFIISPEVIRKFIILVVPPLPACINPLLYLLENPHFRDRLRLRPRA-----GDSGPLAYAA 521
OY 522 AGELEKSSCDSTQALVAFSDVDLLEASEAGRP-----PGLFTYGPSPSYTLISC 570
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 640 SDOVEKQSCDSTQALVFTSSSITVDLPSSSVPSPAYVTESCHLSVAFAVPC 992

```

RESULT 14

ID AAY90687 standard; Protein: 907 AA.

```

XX AAY90687;
AC
XX 21-AUG-2000 (first entry)
DT
XX Human mutant G protein-coupled receptor HG38 (V765K).
DE
XX
XX G protein-coupled receptor; GPCR; constitutively active;
XX intracellular loop 3; transmembrane domain 6; drug screening;
XX agonist; antagonist; mutant; mutelin.
XX

```

OS Homo sapiens.
 OS Synthetic.
 PN MO200022129-A1.
 PD 20-APR-2000.
 XX 12-OCT-1999; 99WO-US23938.
 XX 13-OCT-1998; 98US-0170496.
 XX (AREN-) ARENA PHARM INC.
 PI Behan DP, Chalmers DT, Liaw CW;
 XX WPI: 2000-329165/28.
 XX N-PSDB; AAA30779.
 DR Non-endogenous constitutively activated human G protein-coupled
 PT receptors, useful for identifying agonists for use as pharmaceutical
 PT agents -
 PS Example 2; Page 332-335; 341pp; English.
 XX The invention relates to constitutively active, non-endogenous versions
 CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
 CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
 CC and AAA30775-A30779). The mutant proteins of the invention contain a
 CC mutation in a portion of the protein comprising the intracellular loop 3
 CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
 CC is substituted for an endogenous residue in IC3 at a position 16 amino
 CC acids N-terminal of an endogenous proline in TM6 to form a sequence
 CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
 CC or Ala, and is preferably Lys. When the endogenous residue at this
 CC position is Lys, this residue is replaced by His. Arg or preferably Ala.
 CC The 15 amino acid stretch between the substituted amino acid and the Pro
 CC may be endogenous, non-endogenous, or a mixture of endogenous and
 CC non-endogenous residues. The constitutively active GPCRs are useful for
 CC identifying antagonists, agonists and partial agonists for use as
 CC pharmaceutical agents. The mutant proteins are also useful in research
 CC settings for elucidating the roles of the receptors in normal and
 CC diseased conditions. Antagonists for a particular GPCR are useful for
 CC treating diseases and disorders associated with that receptor. Because
 CC the novel mutant GPCRs are constitutively active, they can be used
 CC directly for screening of compounds without the need for endogenous
 CC ligands. Sequences AAY90643- AAY90677 and AAY90683-Y90687 the mutant
 CC human GPCRs of the invention.
 CC
 SQ Sequence 907 AA;
 XX
 Query Match 38.4%; Score 1269; DB 21; Length 907;
 Best local Similarity 46.7%; Pred. No. 8.9e-107;
 Matches 249; Conservative 96; Mismatches 158; Indels 30; Gaps 8;
 QY 54 KLEKDTFSQSLSDALDSNNATSTPEAFSTLHSLVKIDTJTNOLTTPLAGIGGLMH 113
 DB EIKVDFTQQLSLFSLNLMANKIAIHHNAFSLPLSLKIDLSNLSLSPFTTGHGLTH 447
 QY 114 LKLGKGNLALSGAFKDESPKRLIEVPAYACCPYGMKASFKAAGOW-----EAEDELH 167
 DB 448 LKLTGNLALOSLISENPELKVEMPRAYCCAFGYENMKYKISNOMKNDSSMDLH 507
 QY 168 LDDESSKRPGLLAROAEHHYDDELDELQLEM-EDSKPHPSVOCSPPTGPKCEYLFE 226
 DB 508 -----KKDAGMFQODE-----RDLEDPLDFEDLKALHSVOCSPSPGPKCEHLLD 556
 QY 227 SWGRLAVALVLLSVLCNGVLTLVFAGGAPLPVPKVVGAINGANTLTIGISGLLAS 286
 DB 557 GMLIRIGWTIAVLTALTCNALVTSTVER-SPLYYSIKLLIGVIAVAMLTGSSAVIAG 615
 QY 287 VDAITFGQFSEYGAAMETGCLGATFLAVLGSSEASVLLTLTAAVQCSVSVCAVAYGKS 346
 DB 616 VDAITFGQFSEYGAAMETGCLGATFLAVLGSSEASVLLTLTAAVQCSVSVCAVAYGKS 675

QY 347 PSLGSVRAGVIGCLALAGLAALPLASVGEYCASPLCLPYAPPEGOPALGFTVALVMNN 406
 DB 676 APFSSLKVIILLCALALMAAVPLLGSKYGASPLCLPL--PFEPSITMGVMAVLILIN 733
 QY 407 SFCFLVAGAVIKIKYCDLPRGDFEAVWDCAMVRHVAWILFADGLLYCPAFYAFSFMGL 466
 DB 734 SLCEPLMTIAVTKLYCNLDKGDLENIDCSMKKHIALTLFTNCILNCPAFYAFSSLIHL 793
 QY 467 FVPTPEAVKSVLLVLLPACLNPLLYLLFNPHFRDRLRLRPRA-----GDSGLPATA 521
 DB 794 FTSPSEVIKFILLVVPDPACLNPLLYLLFNPHFKEDLVSLRKQYVWTRSKHPSLMSIN 853
 QY 522 AGELEKSSCDSTQALVAFSDVDLLEASAGRP-----PGLERYGPPSYTLISC 570
 DB 854 SDVEKOSCDSTQALVFTSSSITTYDLPSSVSPAYPVYESCHLSSVAFPVC 906
 RESULT 15
 AAB95066
 ID AAB95066 standard; Protein; 230 AA.
 AC AAB95066;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:16921.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX
 XX EPI074617-A2.
 PD 07-FEB-2001.
 XX
 PD 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELT-) HELIX RES INST.
 XX
 PI Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS Claim 8; SEQ ID 16921; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a 3'-end sequence, where the
 CC polynucleotide which comprises at least 15 nucleotides and the combination
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 230 AA;

Query Match 36.58; Score 1206; DB 22; Length 230;
 Best Local Similarity 99.68; Pred. No. 7.3e-102;

Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 404 MNNSFCFLVYAGAYIKLYCDLPRGDFEAVWDCAMVRYHAWLIFADGLYCPYAFISFASM 463
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MNNSFCFLVYAGAYIKLYCDLPRGDFEAVWDCAMVRYHAWLIFADGLYCPYAFISFASM 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 464 LGLEPYTPPEAVKSVLLVYLPACLNPLLYLLEFNPFRDRLRLRPRAGDSGLAYAAAG 523
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 LGLEPYTPPEAVKSVLLVYLPACLNPLLYLLEFNPFRDRLRLRPRAGDSGLAYAAAG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 524 ELEKSSCDSTQALVAFSDVDLILEASEAGRPGLETYGFPSVTILISCOOPGAPRLGSHC 583
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 ELEKSSCDSTQALVAFSDVDLILEASEAGRPGLETYGFPSVTILISCOOPGAPRLGSHC 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 584 VEPGHNHFGNPPQPSMDGELLRAEGSTPAGGGLSGGGGFQPSGLAFASHV 633
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 VEPGHNHFGNPPQPSMDGELLRAEGSTPAGGGLSGGGGFQPSGLAFASHV 230

Search completed: November 8, 2002, 19:34:26
 Job time : 37.8291 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:50 ; Search time 36.5581 Seconds
(without alignments)
4148.213 Million cell updates/sec

Title: US-09-851-595-8
Perfect score: 3850
Sequence: 1 GLHNLETTDLNKNKLGEPFV.....GGLSGGGGFGPQSLAFASHV 736

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.virus:*
12: sp.vertibrate:*
13: sp.unclassified:*
14: sp.rvivirus:*
15: sp.bacteriophage:*
16: sp.bacteriophage:*
17: sp.archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3850	100.0	928	4 Q9BYD7	Q9BYD7 homo sapien
2	2087	54.2	1360	11 Q8R301	Q8R301 mus muscula
3	840.5	21.8	1360	5 Q9ND11	Q9ND11 drosophila
4	839.5	21.8	1050	5 Q9BN18	Q9BN18 drosophila
5	827.5	21.5	1300	5 Q9NMD6	Q9NMD6 drosophila
6	766.5	19.9	1012	5 Q9SYT6	Q9SYT6 asterina pe
7	766.5	19.9	1280	5 Q9SYT7	Q9SYT7 asterina pe
8	545.5	14.2	701	13 Q9DCG6	Q9DCG6 oreochromis
9	504.5	13.1	688	11 Q64183	Q64183 rattus sp.
10	490.5	12.7	695	11 Q8R428	Q8R428 cavia porce
11	487	12.6	724	13 Q9PVP0	Q9PVP0 oncorhynch
12	482.5	12.5	658	13 Q9PVP9	Q9PVP9 oncorhynch
13	478.5	12.4	779	13 Q918N7	Q918N7 morone saxa
14	475.5	12.4	693	13 Q9DCG5	Q9DCG5 oreochromis
15	475.5	12.4	829	5 Q9VEG4	Q9VEG4 drosophila
16	475.5	12.4	831	5 Q94979	Q94979 drosophila

17	470	12.2	778	13 Q98TF4	Q98TF4 oreochromis
18	468.5	12.2	831	5 Q8SX01	Q8SX01 drosophila
19	460.5	12.0	763	6 Q9BGN4	Q9BGN4 felis silve
20	459	11.9	696	13 Q9DGF5	Q9DGF5 cynops pyr
21	454.5	11.8	764	11 Q9D697	Q9D697 mus musculu
22	450	11.7	701	6 Q8SP58	Q8SP58 bos taurus
23	442.5	11.5	793	13 Q9BGS5	Q9BGS5 oncorhynch
24	436.5	11.3	764	6 Q8SP59	Q8SP59 sus scrofa
25	436.5	11.3	764	6 Q8SP59	Q8SP59 sus scrofa
26	434	11.3	739	6 Q9PW16	Q9PW16 ciarias gar
27	433	11.2	739	6 Q9BGS5	Q9BGS5 sus scrofa
28	432.5	11.2	601	13 Q42500	Q42500 meleagris g
29	431	11.2	601	13 Q18759	Q18759 caenorhabdi
30	428.5	11.1	929	5 Q18759	Q18759 oncorhynch
31	425	11.0	662	13 Q9T949	Q9T949 ictalurus p
32	420.5	10.9	662	13 Q9T854	Q9T854 ictalurus p
33	420	10.9	696	13 Q15966	Q15966 homo sapien
34	413	10.7	701	4 Q14751	Q14751 homo sapien
35	382	9.9	533	4 Q9P1V4	Q9P1V4 homo sapien
36	371.5	9.6	470	4 Q43200	Q43200 homo sapien
37	366.5	9.5	410	4 Q16225	Q16225 homo sapien
38	338	8.8	673	13 Q90WP8	Q90WP8 podarcis si
39	328	8.5	590	6 Q27986	Q27986 bos taurus
40	311	8.1	293	6 Q95ME7	Q95ME7 canis fami
41	284	7.4	335	6 Q46387	Q46387 mustela vis
42	270.5	7.0	307	6 Q95L59	Q95L59 alluropoda
43	259.5	6.7	228	11 Q9BSJ8	Q9BSJ8 rattus norv
44	255.5	6.6	307	13 Q90XC7	Q90XC7 salmo salar
45	255.5	6.6	332	13 Q8QFN6	Q8QFN6 elaphe quad

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	928 AA.
Q9BYD7			
1	Q9BYD7		
2	01-JUN-2001 (TREMURel. 17, Created)		
3	01-JUN-2001 (TREMURel. 17, Last sequence update)		
4	01-JUN-2002 (TREMURel. 21, Last annotation update)		
5	VTS20631 (Fragment).		
6	VTS20631.		
7	GN		
8	OS Homo sapiens (Human)		
9	OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
10	OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
11	OX NCBI_TaxID=9606;		
12	FN [1]		
13	RP SEQUENCE FROM N.A.		
14	RA Okaze H., Hayashi A., Kozuma S., Saito T.;		
15	RT "A member of g-protein coupled receptor family.";		
16	RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
17	DR EMBL; AB049405; BAB39854.1; -		
18	DR HSSP; P23945; IXON.		
19	DR InterPro; IPR000276; GPCR_Rhodopsn.		
20	DR InterPro; IPR001611; LRR.		
21	DR InterPro; IPR003592; LRR_out.		
22	DR InterPro; IPR003591; LRR_Typ.		
23	DR Pfam; PF00560; LRR_13.		
24	DR PRINTS; PR00237; GPCR_Rhodopsn.		
25	DR PRINTS; PR00019; LRR_RichRP.		
26	DR SMART; SM00370; LRR_9.		
27	DR SMART; SM00369; LRR_Typ; 14.		
28	FT NON_TER		
29	SEQUENCE 928 AA; 100487 MW; 4C3364ADEA89C463 CRC64;		
30	Query Match	100.0%;	Score 3850; DB 4; Length 928;
31	Best Local Similarity	100.0%;	Pred. No. 2.5e-292;
32	Matches 736; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
33	1 GLHNLETTDLNKNKLGEPFVAIRTLGRIQLGPHNNNIKAIPKAFMGAPLLQTHFYDN 60		
34			
35	193 GLHNLETTDLNKNKLGEPFVAIRTLGRIQLGPHNNNIKAIPKAFMGAPLLQTHFYDN 252		


```

QY 61 PLOFVGSARQYLPKLTHTLSLNGAMDIOEPDLKGTSLLETLTLTRAGTRILPSCMCOL 120
DB 253 PLOFVGSARQYLPKLTHTLSLNGAMDIOEPDLKGTSLLETLTLTRAGTRILPSCMCOL 312
QY 121 PRUVRLESLHNOIEELPESLHRCOKLEETIGLOHNRWEIGADTFQSLSLOALDLSMNAIR 180
DB 313 PRUVRLESLHNOIEELPESLHRCOKLEETIGLOHNRWEIGADTFQSLSLOALDLSMNAIR 372
QY 181 SIHPEAFSTHSHVYKDLDTNOLTLPLAGLGGMLHKLKCNALISQAFSKDSFPKRLT 240
DB 373 SIHPEAFSTHSHVYKDLDTNOLTLPLAGLGGMLHKLKCNALISQAFSKDSFPKRLT 432
QY 241 EYVYAYQCCPYGMCASFKAAGQWEADLHLDDESSKRPGLLARQENHYDDDLDELQ 300
DB 433 EYVYAYQCCPYGMCASFKAAGQWEADLHLDDESSKRPGLLARQENHYDDDLDELQ 492
QY 301 LEMEDSKPHSVQCCSPTPGEPKCEYLFESWGRFLAWAYLVSLCNGVLTLTVFAGP 360
DB 493 LEMEDSKPHSVQCCSPTPGEPKCEYLFESWGRFLAWAYLVSLCNGVLTLTVFAGP 552
QY 361 APLPVKEVYVAGTAGANTLTGISCGLASVDALTFQGFSEYGARWETGLGRATGFLAVL 420
DB 553 APLPVKEVYVAGTAGANTLTGISCGLASVDALTFQGFSEYGARWETGLGRATGFLAVL 612
QY 421 GSEASVLLTLTAAVQCSVSVSCYRAYGKSPSLGSVRAGVLGCLALAGLAAALPLASVGEY 480
DB 613 GSEASVLLTLTAAVQCSVSVSCYRAYGKSPSLGSVRAGVLGCLALAGLAAALPLASVGEY 672
QY 481 GASPLCLPAPEEGQPAALGFTVALVMNNSFCFLVYAGATIKLYCDLPKDFEAVNMCAM 540
DB 673 GASPLCLPAPEEGQPAALGFTVALVMNNSFCFLVYAGATIKLYCDLPKDFEAVNMCAM 732
QY 541 VRHVAWLIFADDLILCPVAFSLFASMLGLFPVTPPEAVKSVLVVLPPLACINPLVLYLFN 600
DB 733 VRHVAWLIFADDLILCPVAFSLFASMLGLFPVTPPEAVKSVLVVLPPLACINPLVLYLFN 792
QY 601 PFRFDLRLRLRAGDSGFLAYAAAGELKSSQDSQALVAFSDVLLLEASASGRPGL 660
DB 793 PFRFDLRLRLRAGDSGFLAYAAAGELKSSQDSQALVAFSDVLLLEASASGRPGL 852
QY 661 EYGFPSVTLISCOQPGAPRLGSHCVPEEGNHFGNPPQSMGELLRLRAGSTPAGGGLS 720
DB 853 EYGFPSVTLISCOQPGAPRLGSHCVPEEGNHFGNPPQSMGELLRLRAGSTPAGGGLS 912
QY 721 GGGGPOPSGLAFASHV 736
DB 913 GGGGPOPSGLAFASHV 928

```

RESULT 2

```

Q8R301 PRELIMINARY: PRT: 459 AA.
AC Q8R301;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Hypothetical 47.9 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC026896; AAH26896.1; -.
KW Hypothetical protein.
FT NON-TER
SQ SEQUENCE 459 AA: 47889 MW: F0100BF073E81762 CRC64;

```

Query Match 54.2%; Score 2087; DB 11; Length 459;
Best Local Similarity 87.68; Freq. No. 7, 5e-155;

```

Matches 402; Conservative 20; Mismatches 37; Indels 0; Gaps 0;
QY 278 KRLPLGLARQENHYDDDLDELQEMEDSKPHSVQCCSPTPGEPKCEYLFESWGRFLAV 337
DB 1 KRLPLGLARQENHYDDDLDELQEMEDSKPHSVQCCSPTPGEPKCEYLFESWGRFLAV 60
QY 338 WALVLLSVLCNGVLTLTVFAGPAPLPVYKVGATAGANTLTGISCGLASVDALTFEQ 397
DB 61 WALVLLSVLCNGVLTLTVFAGPAPLPVYKVGATAGANTLTGISCGLASVDALTFEQ 120
QY 398 FSEYGARWETGLGRATGFLAVGSEASVLLTLTAAVQCSVSVSCYRAYGKSPSLGSVRA 457
DB 121 FAEYGARWESGLGCATGFLAVGSEASVLLTLTAAVQCSVSVSCYRAYGKSPSLGSVRA 180
QY 458 GVLGCLALAGLAAALPLASVGEYASPLCLPAPEEGQPAALGFTVALVMNNSFCFLVYA 517
DB 181 GALGCLALAGLAAALPLASVGEYASPLCLPAPEEGQPAALGFTVALVMNNSFCFLVYA 240
QY 518 GAYIKLYCDLPKDFEAVNMCAMVRHVAWLIFADGILYCPVAFSLFASMLGLFPVTPPEAV 577
DB 241 GAYIKLYCDLPKDFEAVNMCAMVRHVAWLIFADGILYCPVAFSLFASMLGLFPVTPPEAV 300
QY 578 KSVLLVVLPLPACINPLVLYLFNPHFRDRLRLRPRAGSGPLATAAGLEKSSQDSQ 637
DB 301 KSVLLVVLPLPACINPLVLYLFNPHFRDRLRLRPRAGSGPLATAAGLEKSSQDSQ 360
QY 638 ALVAFSDVLLLEASASGRPGLTYGFPSPVTLISCOQPGAPRLGSHCVPEEGNHFGN 697
DB 361 ALVAFSDVLLLEASASGRPGLTYGFPSPVTLISCOQPGAPRLGSHCVPEEGNHFGN 420
QY 698 QPSMDGELLRLRAGSTPAGGGLFQPSGLAFASHV 736
DB 421 QPSMDGELLRLRAGSTPAGGGLFQPSGLAFASHV 459

```

RESULT 3

```

Q9NDI1 PRELIMINARY: PRT: 1360 AA.
AC Q9NDI1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE Glycoprotein hormone receptor II.
GN RK OR CG8930.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CANTON S; TISSUE-WHOLE ANIMAL;
RX MEDLINE=20359836; PubMed=10899142;
RA Eriksen K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA Grimelikhuijzen C.J.P.;
RT "Molecular Cloning, Genomic Organization, Developmental Regulation,
RT and a Knock-Out Mutant of a Novel Leu-Rich Repeats-containing G
RT Protein-Coupled Receptor (DLCR-2) from Drosophila melanogaster.";
RL Genome Res. 10:924-938(2000).
DR EMBL: AF142343; AAFF6608.1; -.
DR HSSP: Q57815; ID3Y.
DR Flybase: FBgn0003255; rK.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 14.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR SMART: SM00370; LRR; 2.
DR SMART: SM00369; LRR_TYP; 5.
DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.
KW Receptor.

```


DB 784 AKFTYVFLPLNSCCNFBLYAINTKQFKKDCVFLCKHFEESRYVGGGPGRCGVAATTKR 843

QY 631 SSCSDTQALVAFSDVDL---ILEASEGRPRGLETTFPSVTLISCOOPAPRL 682

DB 844 G-----DLPPFLPAAVAHPPGCR-----CLRMLPMPMWHKME 879

RESULT 5

OS 09NKG6 PRELIMINARY; PRT: 1300 AA.

AC 09NKG6; Q9YUJ3; TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical 144.0 kDa protein (Rk gene product).

GN Rk OR CG8930.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazey R., Davis T., Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D., Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A., Palazzolo M., Reese M.G., Spradling A., Tang G., Wan K., Whitehead K., Celinker S., Rubin G.M., "An exploration of the sequence of a 2.9-Mb region of the genome of Drosophila melanogaster: the Adh region."

RT Genetica 153:179-219(1999).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA Celinker S.E., Aghayani A., Arcaina T.T., Baxter E., Blazey R.G., Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M., Farian D.E., Galle R., George R.A., Harris N.L., Hoskins R.A., Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B., Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeita A., Setti H., Snit E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R., Zieran L.L., Rubin G.M.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA MEDLINE=20196006; PubMed=10731132;

RA Amonatides P.G., Scherer S.E., Holt R.A., Evans C.A., Gocayne J.D., Adams M.D., Celinker S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., April J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster K.G., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K., Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hermandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C., Jaisi M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeter F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Splier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhu H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003408; AAF4846.1;

DR EMBL: AE003642; AAF5367.2;

DR Flybase: FBgn003255; rk.

DR InterPro: IPR000276; GPCR_Rhodopsin.

DR InterPro: IPR001611; LRR.

DR InterPro: IPR003592; LRR_out.

DR InterPro: IPR003591; LRR_typ.

DR Pfam: PF00001; 7tm_1; 1.

DR Pfam: PF00560; LRR_10.

DR PRINTS: PR00237; GPCR_Rhodopsin.

DR SMART: SM00370; LRR.

DR SMART: SM00369; LRR_typ; 4.

DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.

KW Hypothetical protein.

SQ SEQUENCE 1300 AA; 144031 MW; BAB9E39F942FA0B3 CRC64;

Query Match

Best Local Similarity 29.4%; Score 827.5; DB 5; Length 1300;

Matches 224; Conservative 128; Mismatches 296; Indels 115; Gaps

19;

QY 5 LEPTIDLVNKRLOEPVPAIRLGRLOELGFNNNNKAIPEKAFMGNPLQTHHYDNP 64

DB 301 LKVLIEDLDNLSLPEGLSKLSOLOELSTSNLRWINDLEPRS--MQLMDRANPLST 358

QY 65 VGRSAFOYLPKHTLSLNGAMDIOEPDCKGTTSLBITLTRAGIRLPSGMCQOPLRL 124

DB 359 ISGARGSKLKLTLSDVKTLSRSEPELACHALBITLKIDRAGIOEVANLCROPLRK 418

QY 125 VELSHNOIEELPS--LHNCQKLEIGLOHNRIMEGATFQSSLSQLV----- 172

DB 419 SLDLSSNOIEKIQGKPEFNKQINDLLSYNNIKAPDPAFGIPLDLYVSIIRIHS 478

QY 173 DLSMAIRSIHPEASTLSLVLKLDITDNLTPLAGLGLMHLKLGNALSGAFSKD 232

DB 479 DLGEWISTYHKEAFSGFTALDNLGNINPELPESGIRALLHKTNPKLREFPPD 538

QY 233 SFPKRLILEVPAYOC--PYGMCASFVKAS----- 261

DB 539 TFPRIOTLILSAVHCACFLPLVAMSSOKTISOEAVLPSPDAEEDMTLNNSMINTP 598

QY 262 -----GQWAEEDLHDEESSRRPLGLARQAEHNYDPLDE----- 298

DB 599 QMHNLKOLGASMDPETA--INFNEEDLOQTGGQIATSYMEYFEHNDVSGATGYCF 657

QY 299 -----IQLEMEDSKPHSPSCSPRPFPKPEYTFEEMGIRLAWAIVLLSLCNGVL 353

DB 658 GTGLFGSMSTEDPQ--GSVQCPLPFGFPLCADLFDWTLRCGVVVFLLSLGNGTVE 716

QY 354 TYVAGGAPLPKPKFYVGAAGANTLIGSCGLASVDALTFQOFSYGRARWEGJGCR 413

DB 717 VLIC--SRKMDVPRFLVCNLAADFWMGIVGLAIVDATIDEFNFALPQMSVLCQL 775

QY 414 TGPLAVGSEASVLLTLAAVQCSVSCRAAGKSPSLGSAVAGLGT--ALAGIAA 471

DB 776 SGPLAVLSLSL--YTLAVITIERNYALTTHLHKRLSLKXGAGYIMSGWFAFLMAL 833

QY 472 LPLASGEIGASPLCPYAPPEGQPAALGFYVALVMNSFCFLVYVAGATIKLXCDPRD 531

DB 834 MPLVGSYDRKRAVCLPFEFTTG--PASLTYYVISMFINCAFLTLMGCYLKMVMAL-RG- 890

QY 532 FEAVW---DCAMRHVAMLIIFADGLICYPVAFLSFASMLGLFPVTPPEAVKSVLLVPLP 568
 DB 891 -SOANNTDSRIAKKAMLLVFTDFLCWSPFIATFISITAFIQLSLQAKIFVAVPLN 949
 QY 589 ACINPLLYLFNPHFRDDLRLRLPR-----AGDSGLPATAAAGLEKSSODSTQALVAF 642
 DB 950 SCGNPFLYAIMKKOKKCVTLCKHFEESSRYVGGGCGRGAVAKRTKNG----- 998
 QY 643 SDVDL---ILEASEAGRPGLTETGFPSTYLSLSCQCAPBLE 682
 DB 999 ---DLPPPLPAAVAHAPGCR-----CLRMPLSEPMNMKKE 1033
 RESULT 6
 Q95Y16 PRELIMINARY; PRT; 1012 AA.
 AC Q95Y16; 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 21, last annotation update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Glycoprotein hormone receptor.
 GN AGPGR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
 OC NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
 RT "cDNA cloning and functional analysis of a novel member of the
 glycoprotein hormone receptor family from a starfish Asterina
 pectinifera." (May-2001) to the EMBL/GenBank/DBJ databases.
 RT Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AB061862; BAB68209.1; -
 DR EMBL: AB061862; BAB68209.1; -
 DR InterPro: IPR002106; AATRNA_1lgaseII.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00060; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 14.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECIP_FL_1; UNKNOWN_1.
 DR PROSITE: PS00262; G_PROTEIN_RECIP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE. 1012 AA; 112623 MW; 52A70E7A88C46B0A CRC64;
 Query Match 19.9%; Score 766.5; DB 5; Length 1012;
 Best Local Similarity 28.8%; Pred. No. 4e-51;
 Matches 197; Conservative 104; Mismatches 287; Indels 95; Gaps 13;
 QY 5 LETTDLNANKLOEPP-VAIRLGLRLOELGFNNNNKAIPEKAFMGN----- 49
 DB 167 LHHNLNDHNLSEVPTALHLSNLRLIHEHNSIPYVVDHAFENSHLIELLRHNT 226
 QY 50 -----PLDQTHFYDNPIDQFVGRSAFOYLPRKLTLSLNGAMDIOEPDLKGTSL 99
 DB 227 HLSAHAFAGLPNMLLEFLGNSITSIATFAFRNLPALENLVILEVKNLSVFPDLTGTSL 286
 QY 100 EITLIRAGIRLPSGMCQDLRLVLELSHNOIELEPILSRCKLE----- 146
 DB 287 EHLGIRCSIRAIPIAFNCNMGTGLSLNLHNNILEGLPSLSCSSLYKLHLGTNKTSL 346
 QY 147 -----ELGLOHNIMEIGADTFQSLSLQALDLSMNAIRSIHPEAFSLHSIVK 195
 DB 347 GQPSGLHLDYDLQLENDISYIPNAPFSLSLHDTLSNNTIREIDSOAFAPCTSIQY 406
 QY 196 LDTLNDLTLPLAGIGLMLKLNALSOAFSKDSPPKRLILEVPAYOCCFYGMCA 255
 DB 407 LDISNNSFPVLPYAGLOMLKIRTYNEOLEDFPPSELPSTEIATATAPYHCCEYELA 466
 QY 256 SFFKASQGWAEADLHDEESSKRPGLGLARQAEHNYDODLDELQLEMEDS----- 306

DB 467 BEYLS-----LADRPNISETTYWASGVDDYNYMTFIDINSESWIDSJFGGSLI 517
 QY 307 -----KPHSVOCSPPTGPFKPECELEFESMGIRLAWAIVLSYLCGLV 352
 DB 518 GSPYLSGYSKSLVPH-NISCRKPPGPMPCMDLFGSWPLRICVWLVLALITAGNAIVI 576
 QY 353 LTVFAGRAPLPRVKEVVGATAGANTLTGICGGLASVDALTFQSESEYCARWETGIGR 412
 DB 577 EVIYSHTKMDVP-REFLONLAFADFPLGVYLGFLAGVDISTLVFRRFGARWOLSAAGR 635
 QY 413 ATGFLAVLGSSESVLLTLAAVQSVSYSCVRAKSPSLGSVAGVGLCA--LAGIAA 470
 DB 636 LAGFLAVSESEFSTYLSYTLTEREYAIKHALHLEKMKL--PHAITMCRGWIJSVTAA 693
 QY 471 ALPLASGEYGAAPLCLPAPEGCPALGFTVALVMNSFCFLVNAQYIKYCDLPRG 530
 DB 694 VLPLVNSHHRVPCLPF--DVDITVAKVYGSILIIILIAFVIMACYSIYATIG 750
 QY 531 DFEAVMDC--AMVRHVAWLIFADGLICYPVAFLSFASMLGLFPVTPPEAVKSVLLVPL 587
 DB 751 SH--AMNCNDSRVARRMSLLVTFDFACWAPIAFPSLTAAGRLSLDGAKVLTIFVPL 808
 QY 588 PACNPLLYLFNPHFRDDLRL 610
 DB 809 NSCANPFLYTLTKQFKKCKTI 831
 RESULT 7
 Q95Y17 PRELIMINARY; PRT; 1280 AA.
 AC Q95Y17; 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 21, last annotation update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Glycoprotein hormone receptor.
 GN AGPGR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
 OC NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
 RT "cDNA cloning and functional analysis of a novel member of the
 glycoprotein hormone receptor family from a starfish Asterina
 pectinifera." (May-2001) to the EMBL/GenBank/DBJ databases.
 RT Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AB061861; BAB68208.1; -
 DR EMBL: AB061861; BAB68208.1; -
 DR InterPro: IPR002106; AATRNA_1lgaseII.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00060; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 14.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECIP_FL_1; UNKNOWN_1.
 DR PROSITE: PS00262; G_PROTEIN_RECIP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE. 1280 AA; 141700 MW; 0AB0ECC0DD880BA CRC64;
 Query Match 19.9%; Score 766.5; DB 5; Length 1280;
 Best Local Similarity 28.8%; Pred. No. 5.4e-51;
 Matches 197; Conservative 104; Mismatches 287; Indels 95; Gaps 13;
 QY 5 LETTDLNANKLOEPP-VAIRLGLRLOELGFNNNNKAIPEKAFMGN----- 49
 DB 167 LHHNLNDHNLSEVPTALHLSNLRLIHEHNSIPYVVDHAFENSHLIELLRHNT 226
 QY 50 -----PLDQTHFYDNPIDQFVGRSAFOYLPRKLTLSLNGAMDIOEPDLKGTSL 99
 DB 227 HLSAHAFAGLPNMLLEFLGNSITSIATFAFRNLPALENLVILEVKNLSVFPDLTGTSL 286

```

QY 100 EITLTRAGILRLPSGMOQLPRLVLELSHNOIEELPSLHRCCKE-----146
D 287 EHLIERKCSIRAIIPANCMDNTGLTSLNINNLLEGRLPSKSSKTEVHLGINKLISLE
QY 147 -----ELGLQNRIRWEIGADTFQSLSLOALDSWMNIRSHHPEAFSTLHSLYK 195
D 347 GQPSGCHLDLYDQLENDISYIPADAFQSLSHLDLTLSSNNTRIDSOAFEPCTSIQY 406
QY 196 LDIITDQTLPLAGLGLMHLTKGNLALSOAFKDSFPRKILEVEPAVOCCPYGMCA 255
D 407 LDISSNSFPLPTAGLOMLKITYNEQLEDPPEPSELPSITETIATAPYHCCEYIELA 466
QY 256 SFPKASGQMEADLHLDDESSKRLGLLARQENHYDDLDLEQLEMDS-----306
D 467 EELYS-----LADPRNISETTYWASGVDPYNNTPFINESWIDSIFGGSLSI 517
QY 307 -----RPHSVQCSPPRPFPKCEYLFESMGIRLAVAVILLVLCNGILV 352
D 518 GSPYLSGNYSLVPH NISCRPKPFPFMDLFGSMPRIQVWLFELLAIIGNALVI 576
QY 353 LTVFAGGAPRLPVKFGVGAIGANTLTGISCGLASVDALTFCQFSEYGARWETGLCCR 412
D 577 FVIIVSHKMPV RFLICNLAFADFLGYLGLAGVDSTLGVFRFGAKMQLSAGCR 635
QY 413 ATGFLAVGSEASVLLTLAAVCCSVSCVRAVGSPLSGVAVGLCLAA LAGLAA 470
D 636 LAGFLAVFSSEFSITYTLSTLTERFYAIKHAHLEKMKL--PHATVWCFGWFISVTAA 693
QY 471 ALPLASVGEASPLCPYPAPREGOPALGFTVALVMNSFCFLVAVAGATKLYCDLPRG 530
D 694 VLPLVNVSHRVPVCLPF--DVDITVAKYVGSIIILNLIAFVIMACASITLAI QG 750
QY 531 DFEAWDC--AMVRHVAWMLIFADGLYCPVAFSPASMLGEPVPEAVKSVLLVLP 587
D 751 SH--ANNCNDSRYARRMSLIVFTDFACMAPIAFFSLTAPGLKLSLDKAKVLLTVLP 808
QY 588 PACINLPLLYLLEFPHRDLLRL 610
D 809 NSCANFELYTLTKOKRCKDKT 831

RESULT 8
QY 09DCG6 PRELIMINARY; PRT; 701 AA.
AC 09DCG6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
GN Gonadotropin receptor I.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Cichlidae; Oreochromis; Perciformes; Labroidae;
OX NCBI_TaxID=6128;
RN SEQUENCE FROM N.A.
RA Oba Y., Hirai T., Yoshinura Y., Yao Z., Nagahama Y.;
RT "Tilapia gonadotropin receptor I.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041762; BAB16106.1;
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 701 AA; 78453 MW; B3D78465CA56410A CRC64;

```

```

Query Match 14.2% Score 545.5; DB 13; Length 701;
Best Local Similarity 26.8%; Pred. No. 4,7e-34;
Matches 175; Conservative 111; Mismatches 306; Indels 61; Gaps 20;

QY 2 LHNLETLIDINKIQEFPV-AIRTLQRLQELGFHNNN-IRAIPEKAPGNLLOTHTPYD 59
D 59 ISMACCLEVKQQLREIQOQGLTSLLOHMLTISENDLESIGFAFGSLPHLKRILSK 118
QY 60 N-PIQFVRSAPFOYLPRKHTLSINGAMDIOEP-PD-LKGTTSLEILTLTRAGIRLPSGM 116
D 119 NAALRNIGAFVSNLPSELITITKSKHSLFIHPDAFRNNAARLFLITSMGIRIPDES 178
QY 117 CQQLRLVLEPLSN-QIEELPS---LHRC-QKLEIGLQHNIRWEIGADTFQSLSLOA 171
D 179 KIHSTACFLDLQDNSHIKRPANAFRGCTQTFPAERLRLRNQIKEVASDAFNG-TKMR 237
QY 172 LDISSN-AIRSHPEASTLSLYKLDITDQTLPLAGLGLMHLTKGNLALSOAFS 230
D 238 LFIQGNQLHISNPAFVGSSELVLDVSETALTSLPDSITLQIKRLIASAFNKLPEP 297
QY 231 KDSFPRKILIEVPAVOCCPYGMCAFFKASQMEADLHLDDESSKRLGLLARQEN 290
D 298 IQLFTKLHQAKLTPPSHCAF--LMHNRNSRMHS---LQDNPAKNNLHFFREYCSN 350
QY 291 HYDQDLDELQLEMDSKPPSVQCSPTPGFPKCEYLFESMGIRLAVAVILLVLCNGI 350
D 351 -----STNITCSPADDPNCPCDIMASAPRLIIVLILGNVAV 392
QY 351 VLTVPAGGAPRLPVKFGVGAIGANTLTGISCGLASVDALTFCQFSEYGARWETGLG 410
D 393 VLL-VLLSRKRLTVPRFLMCHLAFADLCMGILYIVATVMDLGRGRYNAIDMQLG 451
QY 411 CRAGFLAVGSEASVLLTLAAVCCSVSCVRAVGSPLSGVAVGLCLAA-----465
D 452 CNAAGFFTFVFASELSEVFTLAIYERMTTHALRLDKRLTRH-----ACIIMTIGWI 505
QY 466 -AGLAAALPLASVGEASPLCPYPAPREGOPALGFTVALVMNSFCFLVAVAGATKLY 524
D 506 FSLAALLPTVIGISSYKSVSICLPM--DVESLVSQFVVCLLNLIAFVCGCYLSTY 563
QY 525 CDPRGDFEAW-DCAVRRHVAWMLIFADGLYCPVAFSPASMLGEPVPEAVKSVLLV 583
D 564 LTRFKPSSAAAHADTRVAQRMALVIFDFICAPISFFAISALKLPLTVSDSKLLVL 623
QY 584 VLPLPACINLPLLYLLEFPHRDLLRLRPPAGDSGLAVAAAGEL---EKSSC 633
D 624 FYIINCSNFFLYAFTRFRDRDFLLAARFG-----LFTKRAQIYRTSSSC 671

RESULT 9
QY 064183 PRELIMINARY; PRT; 688 AA.
AC 064183;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
OS Follicle-stimulating hormone receptor.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN SEQUENCE FROM N.A.
RA MEDLINE-92149579; PubMed-1738373;
RA Heckert L.L., Daley I.J., Griswold M.D.;
RT "Structural organization of the follicle-stimulating hormone receptor
gene.";
RL Mol. Endocrinol. 6:70-80(1992).
DR EMBL; S81198; AAB21415.2;
DR EMBL; S81117; AAB21415.2; JOINED.
DR EMBL; S81119; AAB21415.2; JOINED.
DR EMBL; S81121; AAB21415.2; JOINED.

```

DR EMBL: S81171; AAB21415.2; JOINED.
DR EMBL: S81174; AAB21415.2; JOINED.
DR EMBL: S81178; AAB21415.2; JOINED.
DR EMBL: S81183; AAB21415.2; JOINED.
DR EMBL: S81185; AAB21415.2; JOINED.
DR EMBL: S81194; AAB21415.2; JOINED.
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 3.
DR Pfam: PF01462; LRRNT; 1.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
SQ SEQUENCE 688 AA; 77341 MW; 441F0D9E7D01DF18 CRC64;

Query Match 13.1%; Score 504.5; DB 11; Length 688;
Best Local Similarity 25.7%; Pred. No. 7,4e-31;
Matches 186; Conservative 106; Mismatches 283; Indels 149; Gaps 25;

DR 30 ELGFHNNNIKAIPKAFMGNPLLOTIHFDNPI-QFVGRSAFOYLPRKLTHTLSNGAMDIO 88
DB 50 ELRFVLTKLRLVIRPKGSFAGSGDLEKIEISQNDVLEADVFSNLPKLTHERIKKANL- 108
QY 89 EPPDLKGTTSLELTTLTRAGIRLLPSGMCQOLPRLRVLELSHQIEELPSLHRCQKLEI 148
DB 109 -----LYINPEAF-QNLPRLKLLISNTGIRKLPVAVHKIQSIQKV 147
QY 149 --GLOHN-RIMEIGADTFSSLSQALDLSMNAIRSIHPEAF-----STLHS 192
DB 148 LLDIDQININIHVARSEFMLSFEW---LSKNGIEEIHNCANGTODLDELNSDNNNEE 204
QY 193 L-----VKLDLTQDLTTLPLAGLGLMHLKLGKGNLALSOAFSKDSPKRLILE 241
DB 205 LRPDVFGASGPVILDISRKVSHSLPNHGLENLKLLARSTYRKKLPLNDKFTVLMAS 264
QY 242 VRYAYGCCPYGMCASFEGKASQWEADLH-----LDEES 276
DB 265 LHYPSHCAAF-----ANLKRQISELHPICNKSILRQIDDMTQIGDQVSLIDDEPS 316
QY 277 SKRPLGLARQAEHNDQDLDELQLEMEDSKPHPSVOCSPPTGPFKPEYLFESNGIRLA 336
DB 317 ---YKGSDDMYNEPDYDCN---EYVD-----YTCSPKPDAPNCEIDMGYNIILRVL 363
QY 337 WVAIVLTVLCNGLVLTTFVAGGAPRLPYKFEVYGATAGANTLTGISGGLASVDALTFEG 396
DB 364 IMFISLITLITGNTTVL-VLTTSQYKLTVPREFLMCNLAFDLCIGIYLLIASVDIHTKS 422
QY 397 QPSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVO-----CSVSVCAVAYGKSP 450
DB 423 QYHNVAIDMOTGAGCDAAGFTVAFASLSTVTLAILERHNTTTHAMQLEC-----KVQ 477
QY 451 SLGSVAGVYGLCALAGLAADPLASVGEYGASPLCPYAPREGQPALGFTVALVMNS 510
DB 478 LRRHAAVMYLG-WTFAFAALFPFIIGISSYMKVSICLPM--DIDSPSLQTLVMMLLVNLV 534
QY 511 PEFLLVAVAGYIKLYCDLPRDF-EAYWDCAMVRHVAWLITFADGLLYCPVAFLSASMGIL 569
DB 535 LAFVYICGCTTHIYLVNRNPTIYSSSDTKIAKRMATLITFDLCMAPISEFPAISASLKV 594
QY 570 PFVTPAVKSVLLVPLPACPLNPLLYTLFNPFRDRLRLPRAGSGPLAAYAAGELE 629
DB 595 PLITVSKATILVLPFVINSKANPLIATFTKNFRDFIL-----LS 637
QY 630 KSSCDSTQALVAFSDVDLLEASAGRPGLETYGPFSVYLISCQGPAGPRLSESHCEP 689
DB 638 KFGCEMAQOIYRT-----ETSSA-----THNF-HARKSHCS--SAPVTVNSYVLP 681
QY 690 EGNH 693
II

DB 682 L-NH 684
RESULT 10
Q8R428 PRELIMINARY; PRT; 695 AA.
AC Q8R428;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Follicle stimulating hormone receptor.
GN FSHR.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_Taxid=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY; TISSUE=TESTIS;
RA Suzuki O.;
RT "Guinea pig follicle stimulating hormone receptor."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY082514; AAL92577.1; -
KW Receptor.
SQ SEQUENCE 695 AA; 77838 MW; 9A3ECFA19C45227B CRC64;

Query Match 12.7%; Score 490.5; DB 11; Length 695;
Best Local Similarity 25.3%; Pred. No. 9,3e-30;
Matches 166; Conservative 111; Mismatches 286; Indels 93; Gaps 19;

QY 30 ELGFHNNNIKAIPKAFMGNPLLOTIHFDNPI-QFVGRSAFOYLPRKLTHTLSNGAMDIO 88
DB 50 ELRFVLTKLRLVIRPKGSFAGSGDLEKIEISQNDVLEADVFSNLPKLTHERIKKANL- 108
QY 89 EPPDLKGTTSLELTTLTRAGIRLLPSGMCQOLPRLRVLELSHQIEELPSLHRCQKLEI 148
DB 109 -----LYINPEAF-QNLPRLKLLISNTGIRKLPVAVHKIQSIQKV 147
QY 149 --GLOHN-RIMEIGADTFSSLSQALDLSMNAIRSIHPEAF-----TLH 191
DB 148 LLDIDQININIHVARSEFMLSFEW---LSKNGIEEIHNCANGTODLDELNSDNNNEE 204
QY 193 L-----VKLDLTQDLTTLPLAGLGLMHLKLGKGNLALSOAFSKDSPKRLIL 240
DB 205 LRPDVFGASGPVILDISRKVSHSLPNHGLENLKLLARSTYRKKLPLNDKFTVLMAS 264
QY 242 VRYAYGCCPYGMCASFEGKASQWEADLH-----LDEESKRLPLG---LARQAE 289
DB 265 LHYPSHCAAF-----ANLKRQISELHPICNKSILRQIDDMTQIGDQVSLIDDEPS 316
QY 277 SKRPLGLARQAEHNDQDLDELQLEMEDSKPHPSVOCSPPTGPFKPEYLFESNGIRLA 336
DB 317 ---YKGSDDMYNEPDYDCN---EYVD-----YTCSPKPDAPNCEIDMGYNIILRVL 363
QY 337 WVAIVLTVLCNGLVLTTFVAGGAPRLPYKFEVYGATAGANTLTGISGGLASVDALTFEG 396
DB 364 IMFISLITLITGNTTVL-VLTTSQYKLTVPREFLMCNLAFDLCIGIYLLIASVDIHTKS 422
QY 397 QPSEYGARWETGLGCRATGFLAVLGEASVLLTLAAVO-----CSVSVCAVAYGKSP 450
DB 423 QYHNVAIDMOTGAGCDAAGFTVAFASLSTVTLAILERHNTTTHAMQLEC-----KVQ 477
QY 451 SLGSVAGVYGLCALAGLAADPLASVGEYGASPLCPYAPREGQPALGFTVALVMNS 510
DB 478 LRRHAAVMYLG-WTFAFAALFPFIIGISSYMKVSICLPM--DIDSPSLQTLVMMLLVNLV 534
QY 511 PEFLLVAVAGYIKLYCDLPRDF-EAYWDCAMVRHVAWLITFADGLLYCPVAFLSASMGIL 569
DB 535 LAFVYICGCTTHIYLVNRNPTIYSSSDTKIAKRMATLITFDLCMAPISEFPAISASLKV 594
QY 570 PFVTPAVKSVLLVPLPACPLNPLLYTLFNPFRDRLRLPRAGSGPLAAYAAGELE 629
DB 595 PLITVSKATILVLPFVINSKANPLIATFTKNFRDFIL-----LS 637
QY 630 KSSCDSTQALVAFSDVDLLEASAGRPGLETYGPFSVYLISCQGPAGPRLSESHCEP 689
DB 638 KFGCEMAQOIYRT-----ETSSA-----THNF-HARKSHCS--SAPVTVNSYVLP 681
QY 690 EGNH 693
II

RESULT 11

```

O9PVPO PRELIMINARY; PRT: 724 AA.
ID O9PVPO
AC O9PVPO;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Gonadotropin receptor II.
GN SGTII.
OS Oncorhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99423507; PubMed=10491336;
RA Oba Y., Hirai T., Yoshikuni M., Kawachi H., Nagahama Y.;
RT "Cloning, functional characterization, and expression of a
gonadotropin receptor cDNA in the ovary and testis of amago salmon
(Oncorhynchus rhodurus).";
RL Biochem. Biophys. Res. Commun. 263:584-590(1999).
DR EMBL; AB030005; BAA84638.1; -.
DR HSSP; P22888; 1LUT.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECCEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECCEP_FL_2; 1.
KW Receptor.
SQ
SEQUENCE 724 AA; 80494 MW; E999652740332B2E CRC64;

Query Match 12.6%; Score 487; DB 13; Length 724;
Best Local Similarity 24.7%; Pred. No. 1.8e-29;
Matches 165; Conservative 113; Mismatches 262; Indels 128; Gaps 21;

1 GLNLTLDLNYNKLQEPVAVIRTLGRLOELGFHNNNIKAIEKAMGNPLDTIHFYDN 60
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
76 GLRVOQHEIGQS-----VALET---LETLAF--NNLUDLNE-IFIKN-IRSLVH--- 118
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
61 PLOFVGRSAFOYLRLKLTLSINGAMDIQEPDLKGTSTSE-----ILTLTRAGIRLLPS 114
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
119 ----IARRTNNLPKRLRLTSISNT-GITVFPDMTSHISLEPPNNQNVLDI----- 163
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
115 GMCOQLRPLRV-----LELSHQIEELPS-LHRQCKLEETIGLOHNRIMEIG 159
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
164 --CDNLVLTLSIPVNAFVGMTETETAMNLFNGGIRELDYAFNGTKINKLVLKNNR----- 216
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
160 ADFEQLSSIQALDLSNMATRSIHPEAFSTLSLVKLDLTLDLTQLTLPLAGLGGLMHLK 219
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
217 -----NLRVHREAFKGAAGPRILDDVSTAIEFLTPSHGLNSVVELVA 258
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
220 KGNLTALSOAFSKDSFPKRLILEVPAAYQCCPYGMCASFEEKASGOME----- 265
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
259 RTAVGLKRLPFRRLGNLQKRAHLTYNSHC-----ALLTWDTHDSPINEAQHNG 308
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
266 AEDLHLDEESSKRPLGLL-----ARQAEHHYDODELDELQLEMEDSKPHPSV 312
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
309 SRPLTCDDSPSKDRPAGVSSDSTSLAVEIHGTNDESGVDFQYRPELML-NCQTRRPTL 367
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
313 QCSPTPEPFPCPELFESWGIRLAWAIVLLSVLQNGVLLTVFVAGGPRPLDPVKVGVCA 372
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
368 QCPEPADAFNPCEIDIAFSEFLRAVIMFINIILATION-LTVLLTFEFSRCKLVPRFLMCH 426
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
373 IAGANTLTGISCGLASVDATLFGQFSEYGARWENGICGRATGGLAVLISEASVLLITLA 432
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
427 LARADFCIGVLLMTAVNDLHTRGHYSERHADMDRGAGCSAAGFLSVFGEGLSVYTLSTI 486
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
433 AVQ-----CSVSVSCVRAVYGKSPSLGVSRAVYLGCIALAGLAALPLASVGEYASPLCLP 488
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 12

```

O9PVN9 PRELIMINARY; PRT: 658 AA.
ID O9PVN9
AC O9PVN9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Gonadotropin receptor I.
GN SGTI-RI.
OS Oncorhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20025349; PubMed=10558873;
RA Oba Y., Hirai T., Yoshikuni M., Kawachi H., Nagahama Y.;
RT "The quality of fish gonadotropin receptors: cloning and functional
characterization of a second gonadotropin receptor cDNA expressed in
the ovary and testis of amago salmon (Oncorhynchus rhodurus).";
RL Biochem. Biophys. Res. Commun. 265:366-371(1999).
DR EMBL; AB030012; BAA86898.1; -.
DR HSSP; P22888; 1LUT.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECCEP_FL_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECCEP_FL_2; 1.
KW Receptor.
SQ
SEQUENCE 658 AA; 72991 MW; 853A38BEA845480 CRC64;

Query Match 12.5%; Score 482.5; DB 13; Length 658;
Best Local Similarity 26.4%; Pred. No. 3.6e-29;
Matches 161; Conservative 100; Mismatches 285; Indels 63; Gaps 14;

30 ELGFHNNNIKAIEKAMGNPLDTIHFYDN-DIOFVGRSAFOYLRLKLTLSINGAMDI- 87
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
56 DLEFGQTHIRVPRQFAFTLMQDLTAIVLFEENGLMESIGARAFANLPRLTEITTKSKHLV 115
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
88 ----QEPDLKGTSTSEITLTLTRAGIRLLPS-GMCOQLRPLRVLELSHN-QIEELPSIH- 140
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
116 IHHQAF--IDLPRKSHLTICNTGLRVLPNFSRISHSAAMTFLDLDODNHHVIVIPSNVAF 172
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
141 ---RQCKLEETIGLOHNRIMEIGADPESQSLQALDSWN-ATRSIHPEAFSTLSLVKL 196
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
173 LGLTNTIDELRLTKNGSSEVESHAENG-TKHKLTLMGMLQSHHNNNSKGAEGEGFL 231
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
197 DLTNDQTLTLPLAGLGMLKLMKLTALSOAFSKDSFPKRLILEVPAAYQCC----- 249
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
232 DISKRLTSLPSVGLGEVHLSAVSFSLRTLPRLPLFTLRLQGANLTPPSGCCAFHHNR 291
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
250 --PYGMCASFEEKASGOMEAEDLHLDEESSKRPLGLIARQAEHHYDODELDELQLEMEDSK 307
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
292 NRTFRMTSACFKPQAO-----NNLHFFMDFCL 318
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```


QY 308 PBPVQSCPTGPFKPCXEFLEFSGIRLAVMAIVLISVLCNGILVLTVPAGPAPLPVK 367
 319 NMTSVACSPAPDAPNCDIMGSAFLRLIWIISVLLALGNITVL-VLLGSRKMTVR 377
 QY 368 FVYVAGNANTLGLSCGLASVDALTFGQSFYGNWETGCGCRATGFLAVGSAVY 427
 378 FLMCHLSFADLMGILVYIVAVDVTRGLYINHAISWOTGACDIAGFTVASELSMF 437
 QY 428 LTLAAVQCSVSVCAVYKSPSGSVRAGVGLCLAGLAALPLASVGEYASPLCL 487
 438 TLTATLLEKHTHTHRLDRKRLRHACAVMATGMAFSCIALALPTVGSSYSKYSICL 497
 QY 488 PYAPPEGAPALGETVALVYMNSECLVYAGAYIKLYCDLPBGDF-EAVYDCAMVRVAN 546
 498 P-MDVESTPSQV-FVAFLLLNVAFLCYVCYCLSTYLSVRNSSPPASAEQMAQMAI 555
 QY 547 LTFADGLVCPVAFLEFASMLGFPVTPKAVSVLLVPLPACLNPLTLFNPFRD 606
 556 LITPDELCAPISEFALSALKPLITVSDSKLLVLYFYINSCANPFLGICRTFRD 615
 QY 607 LRLRPAG 615
 DB 616 FFLAARYG 624

RESULT 13
 Q918N7 PRELIMINARY; PRT; 779 AA.
 ID 0918N7;
 AC 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Thyrotropin receptor precursor.
 OS Morone saxatilis (Striped bass).
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 NC NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVARY;
 RX MEDLINE=20457109; PubMed=11000515;
 RA Kumar R.S., Jirli S., Kight K., Swanson P., Diltman A., Alok D.,
 RA Zohar Y., Trent J.M.;
 RT Cloning and functional expression of a thyrotropin receptor from the
 RT gonads of a vertebrate (bony fish): potential thyroid-independent role
 RL Mol. Cell. Endocrinol. 167:1-9(2000).
 RL EMBL: AF239761; AAF80596.1; -.
 DR HSSP: P16473; 1XUM.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 2.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; UNKNOWN_1.
 DR PROSITE: PS0262; G_PROTEIN_RECPT_F1_2; 1.
 KW Receptor; Signal.
 FT SIGNAL 1
 FT SIGNAL 22
 FT SIGNAL 22
 SEQUENCE 779 AA: 87400 MW: 8686D121B676A9EF CRC64;

Query Match 12.4%; Score 478.5; DB 13; Length 779;
 Best Local Similarity 23.5%; Pred. No. 9,4e-29;
 Matches 177; Conservative 118; Mismatches 315; Indels 143; Gaps 21;
 QY 3 HNLFTLDLNYKLOEPVAIRLGLQELGFHNNNTKALPERAFKGNPLQITHF-YDNP 61
 DB 38 HISCEDIDI-LPRPASTETLMLFE-----TSLSSVPADAFSSVMNSTRYISVDWT 89
 QY 62 IOFVGSAFOYLPKLTLSINGAMDIOEPFDKGTSTLELTTLTRAGIRLLPSCMQQLP 121

DB 90 LQRLERHSEYSLRKTIHTEIRNA-----KSLTYID-----PRAF-KNLP 127
 QY 122 RLRYVLESHNQITELPSLHRCCK-----LEEIGLOHNRIMWEGADTFEQLS-LOALDLS 175
 DB 128 NLKYLGFNTGLTFEPDLNITSNDNFLEI-VDPHYITEIFNPSRGITSVLTFVMT 186
 QY 176 WNAIRSIHPEAFS-----TLHSYKLDLDTNQLTTLPLAG 210
 DB 187 GNGFRELIOHNAFGNTKLDQVDLHNNKYTKMDERAFAGTISGMLDLVSLGTTLPTTG 246
 QY 211 LGLIMLKLGNLALSAFESKDSFPRKRIEVPAYAGCC-----PYGMC--AS 256
 DB 247 MDSIRELKARNAMALKPLPKTFKHLTIANLTYPRHCCGFKMLKKRGLEYITCLMTA 306
 QY 257 FF-----KASQGEADLH----- 270
 DB 307 FYDOHNRKSRVGLRIPSLQGESVETIPDQEPNDGHRSEQODRRDFFGSLHYHAYFG 366
 QY 271 -LDDE-----ESSKRPGLLARQAEENHYDODDLDELQLEMEDSKPHPSVQSPPTGPKPC 324
 DB 367 GQPDVDVGFGETLKNPQEDTSQDFDSRDYVCE--EGEE-----VSCAPVDEERNPC 417
 QY 325 EYLFESNGIRLAVMAIVLISVLCNGILVLTVPAGPAPLPVKFVVGALAGANTLTGISC 384
 DB 418 EDINGFGLRVSVFVSLAVGNVYLL-VLTSYKLSYRFLMCHLAFADLCMGITYL 476
 QY 385 GLASVDALTFGQSFSEYGARWETGICRATGFLAVLGSEASVLLITLAAVQCSVSVSVR 444
 DB 477 LLIASVDLHTAEYFNHAIIDQGTGCGLAGFTVYASELSVYTLVTLTRWYATFAM 536
 QY 445 AYKSPSISGSRVAGVGLCLAGLAALPLASVGEYASPLCLPYAPPEGAPALGETVA 504
 DB 537 RLDRKLHHAAYVLLGWMIFCLLALPLVGVSSYQVSVICLPR-DTOSVNAVYILS 594
 QY 505 LYMNPSFCFLVAGAYIKLYCDLPBGDEA-VMDCAVRYHAMLTFADGLYCPAYLSF 563
 DB 555 VLVNLIILAFVVICAFYICAVHNPHRRSGSKDPTNIAKRMALVLTFTFLCAPISTFYAM 654
 QY 564 ASMTGLFVTPKAVSVLLVPLPACLNPLTLFNPFRDRLRPAGDSPLVYA 623
 DB 655 SAVIDRLPLITVSNKRTLLVLFPLNSCANPELYLFTAFRGDVTLLSKVLCQORAL 714
 QY 624 AAGELEKSSCDSTOALVAFSDVDLILEASEAGR 656
 DB 715 FRGQIVSSKSGSGSTQVR-RDMKVRKSGSGQ 746

RESULT 14
 Q9DGC5 PRELIMINARY; PRT; 693 AA.
 ID 09DGC5;
 AC 09DGC5;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Gonadotropin receptor II.
 GN TGRH-RIL.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Labroidae;
 OC Cichlidae; Oreochromis.
 NC NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oda Y., Hirai T., Yoshitani Y., Yao Z., Nagahama Y.;
 RT "Tilapia gonadotropin receptor II." to the EMBL/GenBank/DBJ databases.
 RL EMBL: AB041763; BAB16107.1; -.
 DR HSSP: P22888; 1LUT.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 2.


```
Db 384 CKSDSGTKEKSTLNPNMPEDEFGSG-----TDDSATDITPIFFAS-----FD 429
QY 294 QDLDELQLE-----MEDSKPH-----PSVQCSPTGPKPKCEYLEESWGI 333
Db 430 YMADDTMKKGTFHEKILNPGDSSAELCGNFTFRKPNIECYPMNDLNPCEVDMGYQWL 489
QY 334 RLAVNAIYLLSYLNCGLVLLTYFAGGPAPLPVKFVYGAAGANTLTGISCGLLASVDL 393
Db 490 RISWNIYVALAVGNVAVLTVLISIRPESTVPVPRFLMCHLAFADCLGLYLLLVACIDAH 549
QY 394 TFGOFSEYGARMETGICRATGFLAVLGSEASVLLTLTAVOCSVSCVRAYGKSPSLG 453
Db 550 SMGEYFNFAIDMOYIGLCKVAGFLTYFASHLSVFTLVTITIERWLAITQAMYLNRHRIKL 609
QY 454 SVRAGVGLCLALAGLAALPLASVGEYGASPLCLPYAPPEGQPAALGFTYVALYMNNSFCF 513
Db 610 PALIMLGMIYSMLMSSLPFGISNYSSTISICLPENNDVYDTI--YLAILGNGVAF 667
QY 514 LVVAGAYIKLYCDLPRGDPEA---VWDCAMVRHVAWLIFADGLLYCPVAFLSFASMLGL 569
Db 668 SIIAVCYAQIYLSLGRHETQAHONSPGELSVAKKMLLVFTNFACWSPIAFGLTALAG- 726
QY 570 FPVTPKAVKSVLLVVL-PLPACLNPLLYLLFNPHERDDLRRLRPAG 615
Db 727 YPLINVTKSKILLVFYPLNSCADPYLYAILTSQYRODLFTLLSKLG 773
```

Search completed: November 8, 2002, 19:36:17
Job time : 43.5581 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:31:55 ; Search time 10.7524 Seconds
(without alignments)
2839.055 Million cell updates/sec

Title: US-09-851-595-8
Perfect score: 3850
Sequence: 1 GLHNLRLDLYNKLQEPV.....GGLSGGGQFQPSGLAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3818	99.2	828	1	LGR6_HUMAN
2	1844	47.9	907	1	LGR5_HUMAN
3	1780	46.2	907	1	LGR3_MOUSE
4	1409	36.6	951	1	LGR4_RAT
5	1409	36.6	951	1	LGR4_HUMAN
6	526.5	13.7	695	1	FSHR_BOVIN
7	524.5	13.6	695	1	FSHR_SHEEP
8	523	13.6	695	1	FSHR_PIG
9	522	13.6	694	1	FSHR_HORSE
10	521.5	13.5	692	1	FSHR_MOUSE
11	513.5	13.3	692	1	FSHR_RAT
12	502.5	13.1	695	1	FSHR_MACFA
13	498.5	12.9	695	1	FSHR_HUMAN
14	492.5	12.8	687	1	FSHR_EOUAS
15	463	12.0	763	1	FSHR_BOVIN
16	460	11.9	764	1	FSHR_CANFA
17	457.5	11.9	693	1	FSHR_CHICK
18	456.5	11.9	696	1	FSHR_PIG
19	456.5	11.9	764	1	FSHR_SHEEP
20	454.5	11.8	764	1	FSHR_MOUSE
21	449.5	11.7	700	1	FSHR_MOUSE
22	446	11.6	701	1	FSHR_BOVIN
23	443.5	11.5	700	1	FSHR_RAT
24	436	11.3	764	1	FSHR_HUMAN
25	425.5	11.1	764	1	FSHR_HUMAN
26	424.5	11.0	764	1	FSHR_RAT
27	421.5	10.9	737	1	LGR8_MOUSE
28	413	10.7	699	1	FSHR_HUMAN
29	411.5	10.7	676	1	FSHR_CALJA
30	394	10.2	538	1	FSHR_SHEEP
31	389	10.1	757	1	LGR7_HUMAN
32	384	10.0	925	1	GLHR_ANCEL
33	357	9.3	366	1	LSHR_CHICK

34	338	8.8	1115	1	GPCR_LYMST
35	233.5	6.1	536	1	GPB8_HUMAN
36	233	6.1	331	1	PUB_AGRBL
37	228	5.9	582	1	SHO2_HUMAN
38	227.5	5.9	603	1	ALS_MOUSE
39	226	5.9	605	1	ALS_HUMAN
40	226	5.9	605	1	ALS_PAPHA
41	224	5.8	582	1	SHO2_MOUSE
42	222.5	5.8	603	1	ALS_RAT
43	220	5.7	1039	1	YR1L_CAEL
44	219.5	5.7	782	1	CHAO_TRICA
45	218	5.7	567	1	GPV_RAT

ALIGNMENTS

RESULT 1

LGR6_HUMAN STANDARD: PRT: 828 AA.

AC Q9HXB8; Q96K69; 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

GN Leucine-rich repeat-containing G protein-coupled receptor 6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20388592; PubMed=10935549;

RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,

RA van der Spek P.J., van Duin M., Hsuen A.J.W.,

RT "The three subfamilies of leucine-rich repeat-containing G protein-

RT coupled receptors (LGR): identification of LGR6 and LGR7 and the

RT signaling mechanism for LGR7."

RL Mol. Endocrinol. 14:1257-1271(2000).

RN [2]

RP SEQUENCE OF 406-828 FROM N.A.

RA Takeda S., Kadowaki S., Haga T., Takasuo H., Mitaku S.,

RT "Identification of G protein-coupled receptor genes from the human

RT genome sequence."

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 431-828 FROM N.A.

RC TISSUE=Mammary gland;

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,

RA Tanase T., Nomura Y., Togiya S., Komai F., Saito K., Yamamoto J.,

RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,

RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.,

RT "NBD human cDNA sequencing project."

RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Orphan receptor.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: AF190501; AAC17168.1; -

DR EMBL: AB083616; BAB89329.1; -

DR EMBL: AK027377; BAB55071.1; ALT_INIT.

DR MIM: 606653; -

DR InterPro: IPR000276; GPCR_Rhodpsn.

```

DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm.1.1.
DR Pfam; PF00560; LRR; 8.
DR PRINTS; PRO1561; EDG8RECEPTOR.
DR PRINTS; PRO0237; GPCRHHODOPS.
DR PRINTS; PRO0019; LEURICHRPT.
DR SMART; SM00370; LRR; 5.
DR SMART; SM00369; LRR_TYP; 10.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECPT_FL_2; FALSE_NEG.
DR C-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
Leucine-rich repeat.
KW DOMAIN 1 428
FT TRANSMEM 429 449
FT DOMAIN 450 459
FT TRANSMEM 460 480
FT DOMAIN 481 505
FT TRANSMEM 506 526
FT DOMAIN 527 548
FT TRANSMEM 549 569
FT DOMAIN 570 588
FT TRANSMEM 589 609
FT DOMAIN 610 635
FT TRANSMEM 636 656
FT DOMAIN 657 670
FT TRANSMEM 671 691
FT DOMAIN 692 804
FT TRANSMEM 813 816
FT REPEAT 22 45
FT REPEAT 46 69
FT REPEAT 71 93
FT REPEAT 94 117
FT REPEAT 118 140
FT REPEAT 142 164
FT REPEAT 188 211
FT REPEAT 212 236
FT REPEAT 238 257
FT REPEAT 258 281
FT REPEAT 283 305
FT DISULFID 503 578
FT CARBOHYD 15 15
FT CARBOHYD 34 34
FT CONFLICT 406 410
FT CONFLICT 628 628
FT CONFLICT 824 824
SQ SEQUENCE 828 AA; 89301 MW; 1B5971445AA2D8B4 CRC64;

Query Match 99.2%; Score 3818; DB 1; Length 828;
Best Local Similarity 99.3%; Pred. 3.2e-265;
Matches 730; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 302 EMDSKPHPSVOCSPPTGPEKPECEYLFESWGIRLAWAIVLLSYLNGVLTTFAGGPA 361
DB 394 EMDSKPHPSVOCSPPTGPEKPECEYLFESWGIRLAWAIVLLSYLNGVLTTFAGGPA 453
QY 362 PLPYKFEVVGATAGATLIGISGLASVDALTFGQFSEYGARWETGLGCRATGFLAVLG 421
DB 454 PLPYKFEVVGATAGATLIGISGLASVDALTFGQFSEYGARWETGLGCRATGFLAVLG 513
QY 422 SEASVLLTLAAVOCSSVSVCAVAYGKSPSLGSRVAGVGLAGLAAALPLASVGEYG 481
DB 514 SEASVLLTLAAVOCSSVSVCAVAYGKSPSLGSRVAGVGLAGLAAALPLASVGEYG 573
QY 482 ASPLCLPAPRPGGQPAALGFTVALVWMSFCPLVAGAYIKLYCDLPRGDFEAWDCAMV 541
DB 574 ASPLCLPAPRPGGQPAALGFTVALVWMSFCPLVAGAYIKLYCDLPRGDFEAWDCAMV 633
QY 542 RHVAMLIADGGLCYPAFLSPASMLGLFPVPEAKSVLTVLPACINPLLYLTFNP 601
DB 634 RHVAMLIADGGLCYPAFLSPASMLGLFPVPEAKSVLTVLPACINPLLYLTFNP 693
QY 602 HFRDDLRLRPRAAGSGPLAAYAGELKSSCDSTQALVAFSDVDLILBSEAGRPGL 661
DB 694 HFRDDLRLRPRAAGSGPLAAYAGELKSSCDSTQALVAFSDVDLILBSEAGRPGL 753
QY 662 TYGFPSTVLISQOGAPRLGSHGVEPEGNHFGNPQPSMDGELLIRAGSTPAGGLSG 721
DB 754 TYGFPSTVLISQOGAPRLGSHGVEPEGNHFGNPQPSMDGELLIRAGSTPAGGLSG 813
QY 722 GGGFQPSGLAFASHV 736
DB 814 GGGFQPSGLAFASHV 828

RESULT 2
ID LGR5_HUMAN STANDARD; PRT; 907 AA.
AC 075473; Q9UP75;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
DE (orphan G protein-coupled receptor HG38) (G protein-coupled receptor
DE 49).
GN GPR49 OR LGR5 OR GPR67.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98308104; PubMed=964214;
RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T.,
RA Liu O.;
RT *Identification and cloning of an orphan G protein-coupled receptor of
RT the glycoprotein hormone receptor subfamily.";
RL Biochem. Biophys. Res. Commun. 247:266-270(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RA *Characterization of two LGR genes homologous to gonadotropin and
RA thyrotropin receptors with extracellular leucine-rich repeats and a G
RA protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
RN [3]
RP FUNCTION: Orphan receptor. It may be an important receptor for
RP signals controlling growth and differentiation of specific
RP embryonic tissues (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal
CC cord, and various region of brain.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).

```


RA Hermey G., Methner A., Schaller H.C., Hermans-Borgmeyer I.;
 RT "Identification of a novel seven-transmembrane receptor with homology
 RT to glycoprotein receptors and its expression in the adult and
 RT Biochem. Biophys. Res. Commun. 254:273-279(1999).
 CC -1- FUNCTION: Orphan receptor. It may be an important receptor for
 CC signals controlling growth and differentiation of specific
 CC embryonic tissues.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC and in the brain. In the central nervous system expression is
 CC restricted to the olfactory bulb. In the adrenal gland detected
 CC only in the neutral-crest derived chromaffin cells of the
 CC medulla, but not in the cells of the adrenal cortex. In the
 CC gonads, the expression is high in Graafian follicle, but absent
 CC from primary and secondary follicles.
 CC -1- DEVELOPMENTAL STAGE: Expressed from embryonic day 10.5 (E10.5) in
 CC the developing spinal cord and in the neuroepithelia of the
 CC myel-, met-, mes-, and diencephalon. Expression is transitory and
 CC the pattern changed rapidly.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF110818; AAD14684.1; -
 DR HSSP: P23945; 1XON.
 DR MGD: MGI:1341817; Gpr49.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_NTerm.
 DR InterPro: IPR003592; LRR_Out.
 DR InterPro: IPR003591; LRR_Typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 3.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_Typ; 8.
 DR PROSITE: PS00237; G-PROTEIN_RECP_FL_1; FALSE_NEG.
 DR PROSITE: PS0262; G-PROTEIN_RECP_FL_2; 1.
 DR G-PROTEIN coupled receptor; Signal; Transmembrane; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT SIGNAL 1 21
 FT CHAIN 22 907
 FT DOMAIN 22 561
 FT TRANSSEM 562 582
 FT DOMAIN 583 593
 FT TRANSSEM 594 614
 FT DOMAIN 615 638
 FT TRANSSEM 639 659
 FT DOMAIN 660 682
 FT TRANSSEM 683 703
 FT DOMAIN 704 723
 FT TRANSSEM 724 744
 FT DOMAIN 745 767
 FT TRANSSEM 768 788
 FT DOMAIN 789 802
 FT TRANSSEM 803 823
 FT DOMAIN 824 907
 FT REPEAT 64 88
 FT REPEAT 89 112
 FT REPEAT 113 136
 FT REPEAT 137 160

FT REPEAT 162 184 LRR 5.
 FT REPEAT 186 208 LRR 6.
 FT REPEAT 209 232 LRR 7.
 FT REPEAT 233 256 LRR 8.
 FT REPEAT 257 279 LRR 9.
 FT REPEAT 281 303 LRR 10.
 FT REPEAT 304 327 LRR 11.
 FT REPEAT 328 350 LRR 12.
 FT REPEAT 351 375 LRR 13.
 FT REPEAT 377 396 LRR 14.
 FT REPEAT 397 420 LRR 15.
 FT REPEAT 442 444 LRR 16.
 FT REPEAT 564 585 LRR 17.
 FT CAROHD 63 63 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAROHD 77 77 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAROHD 208 208 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAROHD 792 792 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 907 AA; 99681 MW; 5531676C0AAE253 CRC64;
 Query Match 46.2%; Score 1780; DB 1; Length 907;
 Best Local Similarity 51.4%; Pred. No. 2,5e-121;
 Matches 351; Conservative 121; Mismatches 193; Indels 18; Gaps 8;
 QY 1 GLHNETTDLNNKIOEPVATRTIGLOEGFHHNNIKAIPEKATMGNDLTQTHFYDN 60
 DB 232 GHLSLETDLNNNDEFFTAIKTSLNKLKGFHNNKISIPRAVGNPSLTTHFYDN 291
 QY 61 PIQVGRSAFOYLPKLTLSLNGAMDIOEPDLKTTSLIYLTRAGIRLPSGMOQL 120
 DB 292 PIQVGRSAFOYLPKLTLSLNGAMDIOEPDLKTTSLIYLTRAGIRLPSGMOQL 351
 QY 121 PLRVLELSHNOIELEPSLRHCOXLEIGLONNRWEIGADTFPSQSLQALDSNNAIR 180
 DB 352 PLRVLELSHNOIELEPSLRHCOXLEIGLONNRWEIGADTFPSQSLQALDSNNAIR 411
 QY 181 SIHPAEFTLSHVLKIDLTDLNLTPLAGLGLMLKLGNTLASQASKDFPKLRIT 240
 DB 412 SIHPAEFTLSHVLKIDLTDLNLTPLAGLGLMLKLGNTLASQASKDFPKLRIT 471
 QY 241 EVRYAYOCPPGMCASFRAQSGQEMAEHLHDESSKRPGLILARQAEHNDQDLEIQ 300
 DB 472 EVRYAYOCPPGMCASFRAQSGQEMAEHLHDESSKRPGLILARQAEHNDQDLEIQ 526
 QY 301 LEM-EDSKRHPVSQCPPTGPGPKCEYLFEESGIRLAWAIVLLSVLCGLVLTYPAGG 359
 DB 527 LDFEEDLNLHVSQCSPPGPKCEYLFEESGIRLAWAIVLLSVLCGLVLTYPAGG 585
 QY 360 PAPLPYKVEVGAALAGANTLTGISGLASVDALTFQGFSEYEGAWETGLGRATGFLAV 419
 DB 586 PAPLPYKVEVGAALAGANTLTGISGLASVDALTFQGFSEYEGAWETGLGRATGFLAV 645
 QY 420 LGSASVLTLLAVVQCVSVSCVRAVYKSPSLGSRVAGLCLADLAGLAALPLASVGE 479
 DB 646 LGSASVLTLLAVVQCVSVSCVRAVYKSPSLGSRVAGLCLADLAGLAALPLASVGE 705
 QY 480 YGASPLCLPYAPPEGOPALGFTVALVMMNSFCFLVAGATIKYCDLPGRGFEAVMCA 539
 DB 706 YGASPLCLPYAPPEGOPALGFTVALVMMNSFCFLVAGATIKYCDLPGRGFEAVMCA 763
 QY 540 MVRVAVMILFADGLICPVAFVAFSAVMDLGFVPEAVKSVLLVPLPACINPLLYLIF 599
 DB 764 MVRVAVMILFADGLICPVAFVAFSAVMDLGFVPEAVKSVLLVPLPACINPLLYLIF 823
 QY 600 NPHRDLRLRLRRA-----GDSGLPAAVAAAGLEKSSCDSTQALVAVASVDLT--LEAS 652
 DB 824 NPHRDLRLRLRRA-----GDSGLPAAVAAAGLEKSSCDSTQALVAVASVDLT--LEAS 883
 QY 653 EAGRP--PLETYGFPSTVLTISC 673
 DB 884 EAGRP--PLETYGFPSTVLTISC 906
 RESULT 4

LGRA_RAT STANDARD: PRT: 951 AA.
 AC Q922H4 2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
 GN GPR48 OR LGRA4.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_taxonomy:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=99065210; PubMed=9849958;
 RA Hsu S.-Y., Liang S.-G., Hsueh A.J.W.;
 RT "Characterization of two LGR genes homologous to gonadotropin and
 RT thyrotropin receptors with extracellular leucine-rich repeats and a G
 RT protein-coupled, seven-transmembrane region.";
 RL Mol. Endocrinol. 12:1830-1845(1998).
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: AF061443; AAC77910.1; -;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_tyr.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 15.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_tyr; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; FALSE_NEG.
 DR PROSITE: PS00262; G_PROTEIN_RECP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; signal; Glycoprotein;
 KW Repeat; Leucine-rich repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 951
 FT DOMAIN 25 544
 FT TRANSSEM 545 565
 FT DOMAIN 566 575
 FT TRANSSEM 576 596
 FT DOMAIN 597 619
 FT TRANSSEM 620 640
 FT DOMAIN 641 661
 FT TRANSSEM 662 682
 FT DOMAIN 683 703
 FT TRANSSEM 704 724
 FT DOMAIN 725 756
 FT TRANSSEM 757 777
 FT DOMAIN 778 783
 FT TRANSSEM 784 804
 FT DOMAIN 805 951
 FT REPEAT 55 79
 FT REPEAT 81 103
 FT REPEAT 104 127
 LRR 1.
 LRR 2.
 LRR 3.

FT REPEAT 128 151 LRR 4.
 FT REPEAT 153 175 LRR 5.
 FT REPEAT 176 199 LRR 6.
 FT REPEAT 200 223 LRR 7.
 FT REPEAT 225 247 LRR 8.
 FT REPEAT 248 270 LRR 9.
 FT REPEAT 272 294 LRR 10.
 FT REPEAT 318 341 LRR 11.
 FT REPEAT 342 366 LRR 12.
 FT REPEAT 368 387 LRR 13.
 FT REPEAT 388 411 LRR 14.
 FT REPEAT 413 435 LRR 15.
 FT DISULFID 618 693 BY SIMILARITY.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 951 AA; 104138 MW; EDD56AC072123461 CRC64;

Query Match 36.6%; Score 1410; DB 1; Length 951;

Best Local Similarity 43.2%; Pred. No. 1.8e-94; Matches 303; Conservative 101; Mismatches 245; Indels 52; Gaps 5;

QY 1 GLNLETLDNLYKLEQFPVIAITLGRLOELGFRNNIKRAIPEKAFGNPLLOTIHYDN 60
 DB 223 GDNLEFLDNTLVNLYLEFPQAIKALPSLKELEGHSNISVYPDAGFGNPLTIHLN 282
 QY 61 PLOFVGRSAFOYLPKLTLLSLNGAMDIOEFDPDLKGTSLLETLTRAGIRLLPSGMCOQL 120
 DB 283 PLFVGSAFAHNSDLHCLVIRGASLVQWEPNLTGVHLSLTITGKISSIPDLCONQ 342
 QY 121 PRLVELSHNOIEELPSLRQCKLEBIGLOHNRITWIGADTESQLSSLOALDSMAIR 180
 DB 343 KMLRTLDLSYNNIRDLPSFNGCRALIEISLRNQDISLKEKTEGTLRLDLISRLIR 402
 QY 181 SIHPEASTLSLVKIDLTNDNLTPLLAGIGLMHLKGNLALSOAFSNDSPFKRLIL 240
 DB 403 EIHSGAFKAGLTITNDLVSNFELSPPTGELINGLNQIKLVGNFKKALARDANRSL 462
 QY 241 EVPAAYOCPEYMGCAFSPFKASGQWEADLHIDDESSRRPLGLIARQENHYDODLELO 300
 DB 463 SVPAAYOCCAFSGWGSYANLNTEDNSPOEHSVTFKEKGTADANVTSAENEHSOI 518
 QY 301 LEMEDSKPHDSVQCPPEPKPCETLFPESWIGLAVAYILSVLCNGVLLVFPAGP 360
 DB 519 -----IIHCTPSVGAFAKPCPEYILGSMWIRLITWFIPLVALNLNLLVLTVEFA-SC 567
 QY 361 APLPEKVVGAIGAMNTITGICGLASVDAITFGOFSEYGAHWETGLGCRATFLAVL 420
 DB 568 SLPASKFLIGLISVSNILMGITGILFLDAVSMGRAPRAEIGIMWETGSGCVAGSLAVE 627
 QY 421 GSEASVLLTLAAVQCSVSVSCVRAVYGRKPSLSGVRAGVLCGLALAGLAALPLASGEY 480
 DB 628 SSESAAVFLTLAAVERSVFAKDLMKHGKSSHLRQFOVALALLALGAAGAAGCFPEHGOY 687
 QY 481 GASPLCLPYAPPEQAPALGFTVALVMNNSFCPLVYAGAYIKIVDCIDRGPDEAVADCM 540
 DB 688 SASPLCLPFP-PTGETPSLGTFTVLNLSLAFLMALIYTKLVCKNLKEDISENSSOSV 745
 QY 541 VRHVAMLFADGLLYCPAFLSFASMLGLFVTEPAVKSVLVYLPACINPLLYLLEN 600
 DB 746 IKHVAMLFTFCIFPCPAFLSFAPLIRAISSIPKIMSVTLIFPPLACINPVLVYFEN 805
 QY 601 PHFRDRLRLRPAG-----DSGPLAVYAAAGLEKSSCDSTQ 637
 DB 806 PKFKEDWKLKRRVTRKHGSVSVSSISOGGCEODEFYDDCGWYSHLONTLVYDCCESEFL 865
 QY 638 ALVAFSDVLLIASEAGRPQLETGFPSVTLISCOQGA 678
 DB 866 LTKPVSKHLI-----KSHSCPVLTAAACQRPBA 894

```

RESULT 5
LGR4_HUMAN
ID LGR4_HUMAN STANDARD; PRT; 951 AA.
AC Q9BX1; Q9NYD1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
   (G protein-coupled receptor 48).
GN GPR48 OR LGR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=21294803; PubMed=11401528;
RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
RT "Molecular characterization of a novel glycoprotein hormone
   G-protein-coupled receptor.";
RL Biochem. Biophys. Res. Commun. 282:757-764(2001).
CC -!- FUNCTION: Orphan receptor.
CC -!- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
   placenta, ovary, testis and adrenal. Expressed also in spinal
   cord, thymoid, stomach, trachea, heart, pancreas, kidney, prostate
   and spleen.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF346711; AAK31153.1;
DR EMBL: AF346709; AAK31153.1; JOINED.
DR EMBL: AF346710; AAK31153.1; JOINED.
DR EMBL: AF257182; AAF68989.1;
DR Genem: HGNC:13299; GPR48.
DR MIM: 606666;
DR InterPro: IPR00276; GPCR_Rhodpsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_out.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 15.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PRO0237; GPCR_RHODOPS.
DR PRINTS: PRO0019; LEURICHRRP.
DR SMART: SM00370; LRR; 6.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_TYR; 15.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor; signal; Transmembrane; Glycoprotein;
KW Repeat: Leucine-rich repeat.
FT SIGNAL 1 24
FT CHAIN 25 951
FT DOMAIN 25 544
FT TRANSSEM 545 565
FT DOMAIN 566 575
FT TRANSSEM 576 596
FT DOMAIN 597 620
FT TRANSSEM 621 641
FT DOMAIN 642 661

```

```

FT TRANSSEM 662 682
FT DOMAIN 683 703
FT TRANSSEM 704 724
FT DOMAIN 725 756
FT TRANSSEM 757 777
FT DOMAIN 778 797
FT TRANSSEM 798 804
FT DOMAIN 805 951
FT REPEAT 55 79
FT REPEAT 81 103
FT REPEAT 104 127
FT REPEAT 128 151
FT REPEAT 153 175
FT REPEAT 176 199
FT REPEAT 201 223
FT REPEAT 224 247
FT REPEAT 248 270
FT REPEAT 272 294
FT REPEAT 318 341
FT REPEAT 342 366
FT REPEAT 368 397
FT REPEAT 388 411
FT REPEAT 413 435
FT DISULFD 618 693
FT CARBOHYD 68 68
FT CARBOHYD 199 199
FT CARBOHYD 294 294
FT CARBOHYD 314 314
FT CARBOHYD 505 505
FT CONFLICT 292 292
FT CONFLICT 433 433
FT CONFLICT 668 668
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFC22CA1BB CXC64;

Query Match 36.6%; Score 1409; DB 1; Length 951;
Best Local Similarity 42.7%; Pred. No. 2,1e-94;
Matches 300; Conservative 106; Mismatches 237; Indels 60; Gaps 9;

```

```

QY 1 GLHNETDLDLVNKLQEPFVATRTIGRLQELGFHNHNNIKATPEKAMGNPLLOTIFYDN 60
DB 223 GLDNETDLDLSNNLGEFFQAIKARPSLKEIGFHSISIVIPDGADGNPLRTTHLYDN 282
QY 61 PIQVGRSAFOYLPRKLTLSLNGANDIOEPDLKGTSLTITLFRAGIRLLPSGMCQOL 120
DB 283 PLSEVGNASAFHNLSDLSLVIRGASWQGFNLTGVHLESILITGTCKSSIPNNLCQDQ 342
QY 121 PLRLVLESHNOIEPLSLRCQKLEITGLQHNRTWEIGADTFSSQLSLQALDLSWNAIR 180
DB 343 KMLRTLDLSYNNIRDLPSFNGCHALEEISLQRNOIYLQIEGTFQGLISRLIDLSRNLIH 402
QY 181 SIHPEAFSTLSLVKLDLTDLQTLTPLAGGIMHKKLGMALISQANSKSPFKRLTL 240
DB 403 EIHSRAPATLPTINLDVSFNEILSFPEGLNGLNQLVGFNKKEALADFNLRSL 462
QY 241 EYPAAYQCCPYGMCASFPEKASQWEADLHLD-----ESSKRPLGLLARAENHYDOL 296
DB 463 SVPAAYQCCAFWGCDSY-----ANLNTEDNSLQDHSVAQEKGTADANANTSTLENESHI 518
QY 297 DELQLEHEDSKPHRSVQCSPTPGPKPCETLFEESGIRALAWAIVLLSVLNGVLLVFE 356
DB 519 -----THCTPSTGAKEPCEYLLGSMIRLITWEIFLVALFMVLITLTF 564
QY 357 AGGPAPLPVYFVVGALAGANTLTGISCGLASVDALTFGGFSEYGANWGLGRATGF 416
DB 565 ASCIS-LPSSKLTGILSVLSNLMFGITGLTFLDVSWGRFAEGTWBEGSGCKVAGF 623
QY 417 LAVIGSASVLLTLAAVQCSVSCVAVYKSPSLGSRVAGVIGCLALAGIAAALPLAS 476
DB 624 LAVFSSSAIFLLMALAVERSLSAKDIKKNKSNHLKQFVALALLAFLGATVAGGFPFLH 683
QY 477 VGEFGASPLCLPYAPRPGQPAALGFYVALVYMNSEFCLVAVGATIKLYCDLPRGDFEAVW 536
DB 684 RGEYSASPLCLP--PGEFPSLGEFTVTLVNLISLAFLMAVIYTKVCNLEKEDSLSENS 741

```

FT	TRANSSEM	444	465	3 (POTENTIAL).
FT	DOMAIN	466	485	3 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	486	508	4 (POTENTIAL).
FT	DOMAIN	509	528	4 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	529	550	5 (POTENTIAL).
FT	DOMAIN	551	573	5 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	574	597	6 (POTENTIAL).
FT	DOMAIN	598	608	6 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	609	630	7 (POTENTIAL).
FT	DOMAIN	631	695	7 CYTOPLASMIC (POTENTIAL).
FT	REPEAT	44	68	LRR 1.
FT	REPEAT	69	93	LRR 2.
FT	REPEAT	119	143	LRR 3.
FT	REPEAT	170	192	LRR 4.
FT	REPEAT	193	216	LRR 5.
FT	REPEAT	218	240	LRR 6.
FT	DISULFID	442	517	BY SIMILARITY.
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	293	293	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	695 AA;	78084 MW;	1EF9DFEFC046380D CXC64;
Query Match 13.7%; Score 526.5; DB 1; Length 695;				
Best Local Similarity 26.1%; Pred. No. 1.5e-30;				
Matches 170; Conservative 108; Mismatches 290; Indels 83; Gaps 15;				

ID FSHR-SHEEP STANDARD; PRT; 695 AA.
 AC P35379; Q28573; Q28574; Q9TS19;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle stimulating hormone receptor).
 GN FSHR.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R1).
 RC TISSUE=Testis;
 RX MEDLINE=93351750; PubMed=8394255;
 RA Yarey T.A., Sairam M.R., Khan H., Ravindranath N., Payne S., Seidah N.G.;
 RT "Molecular cloning and expression of the ovine testicular follicle stimulating hormone receptor.";
 RT Mol. Cell. Endocrinol. 93:219-226(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R2) AND FSH-R3).
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RX MEDLINE=93176195; PubMed=8439338;
 RA Khan H., Yarey T.A., Sairam M.R.;
 RT "Cloning of alternatively spliced mRNA transcripts coding for variants of ovine testicular follicle stimulating hormone receptor lacking the G protein coupling domains.";
 RT Biochem. Biophys. Res. Commun. 190:888-894(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
 RX MEDLINE=98031015; PubMed=9364440;
 RA Yarey T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W.;
 RT "Molecular cloning, structure, and expression of a testicular follicle stimulating hormone receptor with selective alteration in the carboxy terminus that affects signaling function.";
 RT Mol. Reprod. Dev. 48:458-470(1997).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
 RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Ovary;
 RX MEDLINE=20391225; PubMed=10527866;
 RA Babu P.S., Jiang L., Sairam A.M., Youyz R.M., Sairam M.R.;
 RT "Structural features and expression of an alternatively spliced growth factor type I receptor for follicle stimulating hormone in the developing ovary.";
 RT Mol. Cell Biol. Res. Commun. 2:21-27(1999).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=98031017; PubMed=9364442;
 RA Sairam M.R., Subbarayan V.S.R.;
 RT "Characterization of the 5' flanking region and potential control elements of the ovine follicle stimulating hormone receptor gene.";
 RT Mol. Reprod. Dev. 48:480-487(1997).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity of isoform FSH-R1 is mediated by G proteins which activate adenylate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but this does not result in activation of adenylate cyclase. Isoform FSH-R3 may be involved in calcium signaling.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; FSH-R1 (shown here), FSH-R2, FSH-R3 and FSH-R4; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and testis, but not in kidney.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/FSH SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.1sb-sib.ch/announce/](http://www.1sb-sib.ch/announce/or send an email to 1license@1sb-sib.ch)
 CC or send an email to 1license@1sb-sib.ch).
 CC -----
 DR EMBL: L07302; AAA31525.1; -
 DR EMBL: L12766; AAA31523.1; -
 DR EMBL: L12767; AAA31524.1; -
 DR EMBL: L36115; AAK70667.1; -
 DR EMBL: AJ131735; CA010495.1; -
 DR EMBL: AF090438; AAC61749.1; -
 DR PIR: JG1493; JG1493.
 DR HSSP: P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 4.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
 FT SIGNAL 1 17
 FT CHAIN 1 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 630
 FT DOMAIN 631 695
 FT REPEAT 4 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT VARSPLIC 126 133
 FT VARSPLIC 135 695
 FT VARSPLIC 224 259
 FT VARSPLIC 260 695
 FT VARSPLIC 643 670
 FT VARSPLIC 671 695
 FT VARSPLIC 695 AA; 78237 MW; FBF75D89D8C0D4B CXC64;
 SQ SEQUENCE
 QY 30 ELGFHNHNKAIPEKAPMGNDPLQTIHFYDNPV-QVGRSAFOYLPKLTSLNCAMDIQ 88
 Db 50 ELRVLTKLRIVRPGARSGFDLEKIEISQNDVLEIVANFSLPKLHEIRIEKANNML 109
 Query Match 13.6%; Score 524.5; DB 1; Length 695;
 Best Local Similarity 26.4%; Pred. No. 2,1e-30;
 Matches 175; Conservative 101; Mismatches 273; Indels 113; Gaps 18;

```

QY 89 EF-PDLKGTSLLEILTRAGIRLLPSCMCOOLPRLRVLELSHNOIEELPSLHRCQLEE 147
   || || || || || || || || || || || || || || || || || || || || ||
DB 110 YIDPD-----AFONLPRLKRLILSNITKIKHLPANVAKIOSLOK 146
   || || || || || || || || || || || || || || || || || || || || ||
QY 148 I--GLQHN-RIMEIGADTFSSLS-SLQALDLSWNAIRSIHPEAF-----STL 190
   || || || || || || || || || || || || || || || || || || || || ||
DB 147 VLDDIODNINIHVTRNSFMGISFESMIMVLSKNGIOETHNCAFNGTQLODELNLSNRL 206
   || || || || || || || || || || || || || || || || || || || || ||
QY 191 HSL-----VKLLTDNQTLTLLAGLGGMLHKLKGNLNLQSOAFKSDSPFKRLI 239
   || || || || || || || || || || || || || || || || || || || || ||
DB 207 EELPNDVFOGASGYILLDISTRIRSLSYGLLENLKLRAKTYHLKRLPSLEKVTLYVE 266
   || || || || || || || || || || || || || || || || || || || || ||
QY 240 LEVPAYOCCPYGMCASFPEKASGQWEADLA-----LDEESSKRPGLLARQA 288
   || || || || || || || || || || || || || || || || || || || || ||
DB 267 ASLTPPSHOCAR-----ANWRROTSDLHPTCKNSILROEVDMWTQARGORISLAEDD 318
   || || || || || || || || || || || || || || || || || || || || ||
QY 289 ENHYODDLELOLEMEDS--KRPVSQCSPTGPRKPCYLFESMCIIRLAVAVIYLSVL 346
   || || || || || || || || || || || || || || || || || || || || ||
DB 319 EPSYAKGPDMMYSEPDYDLCSEVVDVTCSPEDAFNPPCEDINGYDILRLVIMFTSLAIT 378
   || || || || || || || || || || || || || || || || || || || || ||
QY 347 CGLVLLTFVAGGAPAPLPVKFVVGAIAGANTLTIGSGLLASVDALTFGQSEYGARWE 406
   || || || || || || || || || || || || || || || || || || || || ||
DB 379 GNILVAV-ILITSQKLYVPRFLMCLAPADLCIGYLLILSVDVHTKSQYHNVAIDMO 437
   || || || || || || || || || || || || || || || || || || || || ||
QY 407 TGLGRATGFLAVLGSSEASVLLTLTAAGV-----CSVSVSC--VRAYGKSPSCSVRA 457
   || || || || || || || || || || || || || || || || || || || || ||
DB 438 TCGAGDAAGFEVTFVASELSVYTLTALTERRMHTITHAMOLECKVHRAASITLVGVV-- 495
   || || || || || || || || || || || || || || || || || || || || ||
QY 458 GVLGCIALLAGLAAALPLAVSGEASPLPLPAPREGOPALGETVALVMNSFCFLVVA 517
   || || || || || || || || || || || || || || || || || || || || ||
DB 496 -----FAFAVALPFIIGISSTYMKVSIPLM--DIDPSQLYVMSLVLANLAVVIC 546
   || || || || || || || || || || || || || || || || || || || || ||
QY 518 GAVIKLYCDLPBGDF-EAVWDCAMVRAVAILFPADGLLYCPVALYFASMLGEPVTPPEA 576
   || || || || || || || || || || || || || || || || || || || || ||
DB 547 GCYTHIYLVNRNNITSSSDPKIAKRAMLIFTDFLCMAPISEFPAISAGLXVPLITVSK 606
   || || || || || || || || || || || || || || || || || || || || ||
QY 577 VNSVLLVLPPLACMLPLLYLFPNPRRDDLRKLRPRAGDSGLATAAGELKSSCDST 636
   || || || || || || || || || || || || || || || || || || || || ||
DB 607 SKILLVFPYPCANPFLYAITFRNFRDFFIL-----LSKFGCYEV 649
   || || || || || || || || || || || || || || || || || || || || ||
QY 637 QA 638
   || || || || || || || || || || || || || || || || || || || || ||
DB 650 QA 651
   || || || || || || || || || || || || || || || || || || || || ||

```

```

RT "Porcine follicle-stimulating hormone receptor";
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC -I- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: L31966; AAA86933.1; -.
DB EMBL: AF025377; AAC24981.1; -.
DB HSSP: P23945; 1XUN.
DB InterPro: IPR000276; GPCR_Rhodpsn.
DB InterPro: IPR001611; LRR.
DB InterPro: IPR000372; LRR_Nterm.
DB Pfam: PR00001; 7tm.1; 1.
DB Pfam: PR00560; LRR; 2.
DB Pfam: PF01462; LRRNT; 1.
DB SMART: SM00013; LRRNT; 1.
DB PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DB PROSITE: PS0262; G_PROTEIN_RECP_FL_2; 1.
DB K G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
DB phosphorylation; Repeat; Leucine-rich repeat.
KW
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT TRANSMEM 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CONFLICT 2 2
FT CONFLICT 13 13
FT CONFLICT 60 60
FT CONFLICT 166 166
FT CONFLICT 215 215
FT CONFLICT 247 247
FT CONFLICT 257 257
FT CONFLICT 334 334
FT CONFLICT 349 349
FT CONFLICT 352 352
FT CONFLICT 383 383
FT CONFLICT 407 407
FT CONFLICT 421 421
FT CONFLICT 427 427

```

```

FT CONFLICT 435 435 D -> N (IN REF. 1).
FT CONFLICT 483 483 L -> V (IN REF. 1).
FT CONFLICT 550 550 T -> I (IN REF. 1).
FT CONFLICT 586 586 A -> V (IN REF. 1).
FT CONFLICT 607 607 S -> L (IN REF. 1).
FT CONFLICT 691 691 R -> H (IN REF. 1).
SQ SEQUENCE 695 AA; 78172 MW; E9EBDB29C79CA50 CRC64;

Query Match
Best Local Similarity 26.1%; Score 523; DB 1; Length 695;
Matches 171; Conservative 103; Mismatches 254; Indels 126; Gaps 17;

30 ELGFHNNNKAIEKAPMGNPLQTHFYDNP1-QVGRSAFOYLPRKHLTSLNGAMDIO 88
DB 50 ELKFVLTAKVIRKGAFFSGDLEKIRISQNDVLEIVANFSLPKLHEIRKANML 109
QY 89 EF-PDLKGTSLLELTFRAGIRLLPFGMCOPLRYLVELSHNOIEELPLHCKOE 147
DB 110 YIDPD-----AFQNLPLKRLISNTGVKHLPAVHKIQLQK 146
QY 148 I--GLOHN-RIMEIGADTFESOLS-SIQALDI-SMNAIRSHPEARS-----TL 190
DB 147 VLLDIDQDNIHHTVERNSFVGLSEFESMILMSKNGIREIHNCARNGTQDELNDL 206
QY 191 HSL-----VKLDLDNQLTTLPLAGLGIMHLKLNALQSARSKDPEPKLRI 239
DB 207 EELPNDVFGASGAPVLDLSRTIRHSLPSYGLNKLRAKSTYLNKLPSLEKEVTLME 266
QY 240 LEVPYAYGCCPYGMCASFPGASQOMEADLH-----LDDE 274
DB 267 ASLTYPSHCCAF-----ANMRROISSLHPCNKSILRQEDVDMQARGQSVLAEDG 318
QY 275 ESKRRRLGLLAQAEHNYDODELEDMEDSKPHPSVOCSPPPGPKCEYEIESWGIR 334
DB 319 ESS-----LAEEFDYMESEFDYDLCNEVD-----VICSPEDDFNPECDIGHDLR 366
QY 335 LAWMAVLLVSLNCGLVLTVEFAGRAPLPYKVVGAIGAMTNGISGGLASVALT 394
DB 367 VLIWFISALAIQGN-IIVVILITTSQKYLTPRFKMCNLAFAIDLCIGIYLLLASVDIHT 425
QY 395 FGQSEYEGARMTGTGACRATGFLAVLGSSEASVLLTLAA-----VOCSSVSC 442
DB 426 KIQYHNYALDMQOTGAGCDAGFEFTVASELSYVTLATLERMHTITHAMQDCKV---- 481
QY 443 VRAYKSPISGVBRGVGLCLALAGLAALPLASVEYGSAPLCPLYPAPREGOPALGFT 502
DB 482 -----QLRNASIMLGMIFALFTVALPPLFGISSYMKVSLCLPM--DIDPSLSQLY 531
QY 503 VALVMNSFCPLVAVAGAYIKLYCDLPRCD-FAVWDCAMVRHVAWLIFADGLYCPAEL 561
DB 532 VSLVLAVLAVFVIGCGYTHILTVRNPRIMSSSDTKAKRAMMLITFDLCMAPISFF 591
QY 562 SFASMLGLFPVTPPEAVKSVLLVPLPACLANPLVLLNPHFEDRLRLRPAG 615
DB 592 AISAALKVPLITVSKSLILVLFYDINSCANPFLAIFTKNFRDVEILLSKRG 645

RESULT 9
FSHR_HORSE STANDARD; PRT; 694 AA.
AC P47799;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicitropin
DE receptor).
GN FSHR.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Testis;
RX MEDLINE=94256960; PubMed=8198575;
RA Robert P., Amsellem S., Christophe S., Benifla J.L., Bellet D.,
RA Koman A., Bidaire J.M.;
RT "Cloning and sequencing of the equine testicular follitropin
RT receptor.";
RL Biochem. Biophys. Res. Commun. 201:201-207(1994).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE. AMONG ALL MAMMALIAN FSH RECEPTORS, ON THE HORSE
CC RECEPTOR DOES NOT BIND LH/CHORIONIC GONADOTROPIN (CG).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSR SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S70150; AAB30854.1; -.
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR00372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 3.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G-PROTEIN_RECPT_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECPT_F1_2; 1.
DR KMW: G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
DR KMW: Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 694
FT DOMAIN 18 365
FT TRANSMEM 366 386
FT DOMAIN 387 397
FT TRANSMEM 398 420
FT DOMAIN 421 442
FT TRANSMEM 443 464
FT DOMAIN 465 484
FT TRANSMEM 485 507
FT DOMAIN 508 527
FT TRANSMEM 528 549
FT DOMAIN 550 572
FT TRANSMEM 573 596
FT DOMAIN 597 607
FT TRANSMEM 608 629
FT DOMAIN 630 694
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 168 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 441 516
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 268 268
FT CARBOHYD 293 293
SQ SEQUENCE 694 AA; 78004 MW; E2F077C5E8BCA54 CRC64;

Query Match
Best Local Similarity 25.6%; Score 522; DB 1; Length 694;
Matches 168; Conservative 113; Mismatches 282; Indels 92; Gaps 15;

30 ELGFHNNNKAIEKAPMGNPLQTHFYDNP1-QVGRSAFOYLPRKHLTSLNGAMDIO 88

```


Db 50 ELRFVLTAKRVIPKGFSGFGLKEIISQNDVLEIVANVSNLEPKLEIRERKANN-- 107
 Oy 89 EPPDLKGTTSLEILITRAGIRLLPSGKCOOLPRVLVLSHNOIELPSLRCKLEI 148
 Db 108 -----LXYIDHAFQNLPLNLYLLSNGIKHLRAVHKIQSLQKV 147
 Oy 149 --GLOHN-RIVEIGADTFQSLS-SIQALDLSMNAIRSIHPEAF-----STLH 191
 Db 148 LLDIQDNINIHVERNSFGLSFESTILRLSKNGIOEINHCAFNQGLDELNTSYNNLE 207
 Oy 192 SL-----VKLDLFDNQLTTLPLAGLGGLMLKLGKLNALQAFSKDSFPKRL 240
 Db 208 ELPNDVFOGASGVIIIDISGRHSIPNGLNKKILRASTYNNLKKPLSEKFAVAMEA 267
 Oy 241 EVRYAVOCPCYGCASFKNAGS-----QWEAEHLHDDESSKRPUG 282
 Db 268 NLVYPHSCAF---AMRRROTSELQTTCKNSILRQEVDMQANGERSVLAEDDESSYPKG 324
 Oy 283 LLARQAEHNYDQDLDELQLEMEDESKRPHSVQCSPTGPEPKPCYELEPSKIRLAVNAIVL 342
 Db 325 FDMWYSFEVD-----LCNEVVD-----VTCSPKPDAPNCEIDIMGYDLRLVLPIS 373
 Oy 343 LSVLCNGVLVITFAGGAPALPRPYKPYVGAAGANTLTGSCGLASVADALTFQCFSEYG 402
 Db 374 LAITGN-IIVLVILITTSQYKLYVPRFLMCNLAFLADLCIGYLLILASVDITHTKQYHNYA 432
 Oy 403 ARMETGCGCRATGFLAVGSEAVLLTLTAAVQCSVSVSCVRAVYKSPSGSVRAVYLCG 462
 Db 433 IDMOTGAGCCDAGFTFYFASLSYTLTATLERMHTTHAMOLECKYQLHNASVMLVG 492
 Oy 463 LALAGLAALPLASVGEYASGAPLCPVAPRPGOPRALGFTVALVMNSFCFLVAVGATIK 522
 Db 493 WIFAFVALLPFIISTYMKVSIQCLPM--DIDSPSLQILVMSLVLANVLAIVIGCIVH 550
 Oy 523 LYCLPRGDF-EAWDCMVRHVAWILFADGLLYCPVAFSLFASMIGLFPVTPAIVKSVL 581
 Db 551 IYLVARNNIYSSSDMYIAKRMALIFTDFLCMAPISEFAISASLAKVPLITVSKILL 610
 Oy 582 LVVLPPLACPLNPLVLYLFPNHRDDLRRLRPACDSGPLAVAAAELEKSSDST 636
 Db 611 VLFVYINSCANPFLYAIPTKFRDFILLSKFG---CYEMQOLYKRTTST 660
 RESULT 10
 FSHR_MOUSE STANDARD; PRT; 692 AA.
 AC P35378; O9QWV8; Q9D4C2;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicle
 receptor).
 GN FSHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv; TISSUE=Testis;
 RA Teana-Sempere M., Manna P.R., Huhtaniemi I.T.;
 RT "Molecular cloning of the mouse follicle stimulating hormone receptor
 complementary deoxyribonucleic acid: functional expression of
 alternatively spliced variants and receptor inactivation by a C566T
 transition in exon 7 of the coding sequence";
 RL Submitted (Sep-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Atakawa T., Hata A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Bash G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=93093308; PubMed=1459341;
 RA Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T.,
 RA Sprengel R.;
 RT "The murine luteinizing hormone and follicle-stimulating hormone
 receptor genes: transcription initiation sites, putative promoter
 sequences and promoter activity.";
 RL Mol. Cell. Endocrinol. 88:55-66(1992).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: SUBFAMILY.
 CC -1- SIMILARITY: CONTRAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF095642; AAC67559.1; -
 DR EMBL: AK016635; BAB30351.1; -
 DR EMBL: S49632; AAB24401.1; -
 DR EMBL: M87570; AAA37641.1; -
 DR MGD: MGI:95583; Fshr.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCRHDOPSN.
 DR SMART: SM0013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 692
 FT DOMAIN 18 365
 FT TRANSMEM 366 386
 FT DOMAIN 387 397
 FT TRANSMEM 398 420
 FT DOMAIN 421 442
 FT TRANSMEM 443 464
 FT DOMAIN 465 484
 FT TRANSMEM 485 507
 FT DOMAIN 508 527
 FT TRANSMEM 528 549
 FT DOMAIN 550 572
 FT DOMAIN

FT TRANSMEM 573 596 6 (POTENTIAL).
 FT DOMAIN 597 607 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 608 629 7 (POTENTIAL).
 FT DOMAIN 630 692 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 44 68 LRR 1.
 FT REPEAT 69 93 LRR 2.
 FT REPEAT 119 143 LRR 3.
 FT REPEAT 170 192 LRR 4.
 FT REPEAT 193 216 LRR 5.
 FT REPEAT 218 240 LRR 6.
 FT DISULFID 441 516 BY SIMILARITY.
 FT CARBOHYD 191 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 436 436 Q -> K (IN REF. 2).
 SQ SEQUENCE 692 AA: 77769 MW: 4857229180563444 CRC64;

Query Match 13.5%; Score 521.5; DR 1; Length 692;
 Best Local Similarity 25.3%; Pred. No. 3,4e-30;
 Matches 186; Conservative 111; Mismatches 272; Indels 165; Gaps 25;

QY 30 ELGFHNNNKAIPEKAFMGNPLLOTIFYDNPI-OEFGRSAPFOYLDPKLHTLSNGAMDIO 88
 DB 50 ELRFVLTAKRVIPKCSFSGFDEKIEISQNDVLEIVLEADVSNLPHNRIKAKNNL- 108
 QY 89 EFPDLKGTSLTLITRAGIRLPSGCOQLPRLVLELSHNOIELDSLARCQLEET 148
 DB 109 -----LYINPEAF-QNLPSELYLLISNGIKHLPAFKIQSLQKV 147
 QY 149 --GLOHN-RIMEIGADTFQSLS-SLOALDSNVAIRSHPEAF-----STH 191
 DB 148 LLDIDNINIHIIANSEFMSGLSFEVITLMLNKGIOEHINCAFNQGLDELNSDNNLE 207
 QY 192 SL-----VKLDLIDNOLTLPLAGLGLMHLKIKMLASQAFSKDSFPLRTL 240
 DB 208 ELPPDFEGAGSPVVDLSRTKRVYSLPHNGLNKKLRARSYRKKLPSLDKFWLIRA 267
 QY 241 EVPAVYQCCPYGMCASFPGKASGOMEAEHLHDESSKRPGLLARQENHNDQDEL- 299
 DB 268 SLTYPSHCCAF-----ANNRROTSELHPLCKNSISR-----QDIDMT 305
 QY 300 -----QLEMESKP-----HPVQCSFPPGPFKCEYLFEFSGWIR 334
 DB 306 QPGDORSVLDVDEPSYSGKSDMLYSEFDYDLCNEFDVTCSPKPAFPCEDIMGNILR 365
 QY 335 LAWAIVLLSLVGLVLTVEFAGGAPAPLPYKPYVGAIAANLTGISCGLASVDALT 394
 DB 366 VLIMFISLITATGNTYLV-VLTISQYKLYVPRFLMCLAFADLCIGYLLDLIASVDIHT 424
 QY 395 FGQSESGARRETGLGCRANGFLAVLGSSEASVLLTLAAV-----QCSYSV 440
 DB 425 KSOYHNVAIMQTAGCDAGFFVFASELISV-YTLAATILERMHTITHAMOLECKYQL 482
 QY 441 SCVARYKSPSLGVSRAVGLGCLALAGLAALPLASVGEYASPLCLPYAPDEGQPALG 500
 DB 483 -CHAA-----SIM-VLG-WAFAPAALEPTEFGISSTYKVSICLPM-DIDPSQL 528
 QY 501 FYVALVMANSEFCLVAVAGATKLYCDLPRGDF-EAVVDCAMVRHVMILPADGLLYCPVA 559
 DB 529 YVALLVYLNALAVVIGCYTHIYLVARNNIYSSSDRTIAKAMALLITFDLCAPIL 588
 QY 560 FLSEFASLGLFPYTPREAVKVVLLVLPRLACNLPILYLLFNPHRDDLRRLRPAGDSG 619
 DB 589 FFAISLSLKVPLTVTSKAKLLLVLPFINSCLAMPFLTAIFTKNFRDFVL----- 639
 QY 620 LAYAAAGELKSSCDSTQALVAFSDVDLLLEASGAPPELETYYGPFSSVLLISQDGP 679
 DB 640 -----MSKECYEVQAOIYKTETSSITTHNHSRKNP-----CS-SAP 675
 QY 680 RLEGSHCVPEEGNH 693
 DB 676 RVTNSYLVPL-NH 688

RESULT 11
 FSHR_RAT
 ID FSHR_RAT STANDARD: PRT: 692 AA.
 AC P20395;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
 GN FSHR
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Seroli cells;
 RX MEDLINE=91125358; PubMed=2126341;
 RA Sprengel R., Braun T., Nikolics K., Segaloff D.L., Seeburg P.H.;
 RT "The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA.";
 RL Mol. Endocrinol. 4:525-530(1990).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92149579; PubMed=1738373;
 RA Heckert L.L., Daley I.J., Griswold M.D.;
 RT "Structural organization of the follicle-stimulating hormone receptor gene.";
 RL Mol. Endocrinol. 6:70-80(1992).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: SEROTOLI CELLS AND OVARIAN GRANULOSA CELLS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: L02842; AAA1175.1; .
 DR PIR: A34548; A34548.
 DR PIR: A41729; A41729.
 DR HSSP: P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR001611; LRR
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PR00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_3.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN RECP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 692
 FT DOMAIN 18 365 FOLLICLE STIMULATING HORMONE RECEPTOR.
 FT TRANSMEM 365 386 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 387 397 1 (POTENTIAL).
 FT TRANSMEM 398 420 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 421 442 2 (POTENTIAL).
 FT TRANSMEM 443 464 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 465 484 3 (POTENTIAL).
 FT TRANSMEM 485 507 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 508 527 4 (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 528 549 5 (POTENTIAL).
 FT DOMAIN 550 572 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 573 596 6 (POTENTIAL).
 FT DOMAIN 597 607 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 608 629 7 (POTENTIAL).
 FT DOMAIN 630 692 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 44 68 LRR 1.
 FT REPEAT 69 93 LRR 2.
 FT REPEAT 119 143 LRR 3.
 FT REPEAT 170 192 LRR 4.
 FT REPEAT 193 216 LRR 5.
 FT REPEAT 218 240 LRR 6.
 FT DISULFID 441 516 BY SIMILARITY.
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 692 AA: 77681 MW: 267547807C9D8EC6 CRC64:

Query Match 13.3% Score 513.5; DB 1; Length 692;
 Best Local Similarity 25.9%; Pred No. 1.3e-29;
 Matches 188; Conservative 106; Mismatches 284; Indels 147; Gaps 25;
 30 ELGFHNNNIKAIEKAFMGNPLQTHFYDNP1-QFVGRSAFOYLPKLTSLGANDIQ 88
 50 ELRFVLTLEKLVIFKSPAGSGDELEKIEISONDVLEADVSNLPKLTHERIKRANVL- 108
 89 EFPDLKGTSLLELTIRAGIRLLPSGMCQPLRLVLELSHNOIEELPSLRHCQKEEI 148
 109 -----LYNPEAF-ONLPSRLILISNTGIRKHLPAVHKOTOSLOKV 147
 149 --GIQHN-RIMEIGADTFPSOLS-SIALDLSWNAIRSIHPEAF-----STLH 191
 148 LLDQDQININIHVARNSFGMLSFESVILMLSKNGIEEIHNCAGFOTDELNLSDNNLE 207
 192 SL-----VKDLTQNLQTLTFLAGGGLMHLKLNLSLQFSKDSPEKRLIL 240
 208 ELPRVDFOGASGPYILDSRTKRVSLPHGLENLKLKARSTYRLKLPILDKVETLMEA 267
 241 EYPAAYOCCPRGMCASFEEKASGOWEADLH-----LDDEE 275
 268 SLTPSHCCAF-----ANLKRQISLHPICKNSILRQDIDMTQIGDQVSLIDEP 319
 276 SSKRPLGLIARQENHNDQDIDLEQLEKEDSKPHPSVCCSPFGFKCEYLFESWGLRL 335
 320 S-----YKGSMDMMNEEFDYDCN--EYVD-----VTCSPKPDFAFNCEDEMGYNILRV 366
 336 AVMALVILSVLCNGVILTVFAGGPAPLPYKVEYGAIALAGANTLTGICGLIASYDALTF 395
 367 LWFSTILATIGNTVLV-VLTTSQYKLTVPREFLMCNLAFLDLCIGIYILLIASVDHTK 425
 396 GOFSEYGARWETGLGCRATGFLAVLGSEASVLLTLAVQ-----CSYSVSCVARYGKS 449
 426 SQYHNYAIDMCGAGCDAAAGFTVFASELSVYTLTAITLERHMTTHHAQLCC-----KV 480
 450 PSLGVRAGVIGCLALAGLAALPLASVGEYASPLCLPYARPEGOPALGFTVALVMAN 509
 481 QLRHNASVWVLD-WTPAPAAALPPIFGISSTYKVSICLPM--DIDSPLSQLYVMALLVN 537
 510 SFCEPLVAVAGATKLYCDLPGRDF-EAVWDCAMVRHVAWILIFADGLIYCYVALISFPMIG 568
 538 VLAFAVYICGCTHYHILTVANPTIVSSSDTKIAKRMATLIFDDELMAIDISFAISASLK 597
 569 LFPVTPPEAVKSVLLVLPPLACINPLLYLLFNPHRBDRLRPRAGDSGPLVAAAGEL 628
 598 VPLITVSKAKILLIVETPNSCANPFLVAFIKNERDRDFIL-----L 640
 629 EKSSCDSTALVAFSDVDLILEASEAGRPGLTETGFPSTYLLISCOOPGAPRLGSHCVC 668
 641 SKFGCEMOQOITRT-----ETSSA-----THNF-HARKSHCS--SAPRTVNSVVLV 684
 689 PEGNH 693
 685 PL-NH 688

RESULT 12
 FSHR_MACFA ID FSHR_MACFA STANDARD; PRT: 695 AA.
 AC P32212;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE FSHR. Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 GN Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_Taxid=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=94071854; PubMed=7504463;
 RX Gomoll J., Dankbar B., Sharma R.S., Nieschlag E.;
 RT "Molecular cloning of the testicular follicle stimulating hormone
 receptor of the non human primate Macaca fascicularis and
 identification of the non human primate Macaca fascicularis and
 Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 RL -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/FSH SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X74454; CAA52463.1; .
 CC PIR: S36452; S36452.
 CC PIR: JN0898; JN0898.
 CC HSSP: P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR00372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECPT_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 KM SIGNAL 1 17
 FT CHAIN 18 695
 FT DOMAIN 18 366
 FT TRANSMEM 367 387
 FT DOMAIN 388 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT TRANSMEM 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT TRANSMEM 598 608
 FT TRANSMEM 609 630

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
 GN receptor).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE-Ovary;
 RP TISSUE-Testis;
 RX MEDLINE=9122171; PubMed=1709010;
 RA Mingesh T., Nakamura K., Takakura Y., Itoh Y., Igarashi M.;
 RL "Cloning and sequencing of human FSH receptor cDNA."
 RN Biochem. Biophys. Res. Commun. 175:1125-1130(1991).
 RP [2]
 RC SEQUENCE FROM N.A.
 RP TISSUE-Testis;
 RX MEDLINE=93246012; PubMed=1301382;
 RA Kelson C.A., Cheng S.V., Nugent N.P., Schweichardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzeja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RL "The cloning of the human follicle stimulating hormone receptor and
 its expression in COS-7, CHO, and Y-1 cells."
 RN Mol. Cell. Endocrinol. 89:141-151(1992).
 RP [3]
 RC SEQUENCE FROM N.A.
 RP Tilly L.T., Alhara T., Nishimori K., Uei X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Hsueh A.J.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RC SEQUENCE OF 1-342 FROM N.A.
 RP TISSUE-Testis;
 RX MEDLINE=93075197; PubMed=1359889;
 RA Gromoll J., Gudermann T., Nieschlag E.;
 RL "Molecular cloning of a truncated isoform of the human follicle
 stimulating hormone receptor."
 RN Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RP [5]
 RC SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=95011044; PubMed=7926278;
 RA Gromoll J., Dankbar B., Gudermann T.;
 RL "Characterization of the 5' flanking region of the human follicle-
 stimulating hormone receptor gene."
 RN Mol. Cell. Endocrinol. 102:93-102(1994).
 RP [6]
 RC 3D-STRUCTURE MODELING OF 49-228.
 RX MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RL "Structural predictions for the ligand-binding region of glycoprotein
 hormone receptors and the nature of hormone-receptor interactions."
 RN Structure 3:1341-1353(1995).
 RP -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 ADENYLYL CYCLASE.
 RP -1- SUBCELLULAR LOCATION: Integral membrane protein.
 RP -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 RP -1- TISSUE SPECIFICITY: SEROTOL CELLS AND OVARIAN GRANULOSA CELLS.
 RP -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 RP FSH/LSH/TSH SUBFAMILY.
 RP -1- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between The Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC NM_156606

OY	192	SL-----VKLDLTNDQLTLLPLAAGLGMHLTKNLAALSOAFSNDSPFKRL	240
		: : : : : : : : : : : : : : : : : : : :	
Db	208	ELPNDFVHGASGVIIIDISRTIRHPSVGLNLEKRLKARSTYNLKKPLTEKVALMEA	267
OY	241	EVPAIVOCPCYVGCASFPAASGOWEMEDH-----LDDEESSKRLGLLARQAE	289
		: : : : : : : : : : : : : : : : : : : :	
Db	268	SLVPSHCAF-----ANMRQISELHPICNKSILRQEVDFWQYRGORSIAEDNE	319
OY	290	NHYDODLDELQLEMEDSKRPP--SVQCSPTPGCFKPEKLEFMSQIRLAWAIVLSVLC	347
		: : : : : : : : : : : : : : : : : : : :	
Db	320	SSYSRGFDMTYTEDYDLCNEVDVYCSKPPAFNCEIDMGVILRLVIMFISLATG	379
OY	348	NGVLVLYFAGGRPLPEPVKEVVGALAGANTLTIGISCGLLASVDALITFGQSEYGARMT	407
		: : : : : : : : : : : : : : : : : : : :	
Db	380	N-IIVLVLTTSQYKLVLPREFLMCNLAFDLGICITLLILASVDITHKSOYHNVAIDWQT	438
OY	408	GLGCRATGFLAVLGSEASVLLTLTAAYQ-----GSVSYGCRANGKSPSLGSRVAGVG	461
		: : : : : : : : : : : : : : : : : : : :	
Db	439	GAGCGAAAFETVFVASELSVYTLTAITLERMHTITHMAQDCC-----KVQLRRAASVWNG	493
OY	462	CLALAGLAALPLASVGEYGASPLCLPYAPPEGCPALGFVTALVMNSFCELVYAGAYI	521
		: : : : : : : : : : : : : : : : : : : :	
Db	494	WTFEPAALPPIRGISYSYMKVSTCLPM--DIDPSQLQVMSLVLVNLVLAIFVLCGYI	550
OY	522	KTYCPLRGDFE-EAYWDCAMVRIHVAWLIFADGLICVYATLSASMLGEPVYPEAKSV	580
		: : : : : : : : : : : : : : : : : : : :	
Db	551	HYLVYRNPNIWSSSDTRIAKRMAMLFTDFDLMAEISFEFASISAKVPLTVSKAKLI	610
OY	581	LVLVPLPACLNPLLYLLENPHFRDRLRLPRAGDSGLPYAAAGLEKSSCDST	636
		: : : : : : : : : : : : : : : : : : : :	
Db	611	LVLEPHINSCANPFLYATFTKFRDFILLSKCG-----CYEQAOIYRTETST	661
		: : : : : : : : : : : : : : : : : : : :	
		RESULT 14	
	FSHR_EQUAS		
ID	FSHR_EQUAS	STANDARD;	PRT; 687 AA.
AC	095179;		
DT	01-NOV-1997 (rel. 35, Created)		
DT	01-NOV-1997 (rel. 35, Last sequence update)		
DT	16-OCT-2001 (rel. 40, Last annotation update)		
DE	Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).		
GN	FSHR.		
OS	Equus asinus (Donkey).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.		
OX	NCBI_Taxid=9793;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	TISSUE=Testis;		
RA	MEDLINE=97338913; PubMed=9195473;		
RT	Richard F., Martial N., Remy J.-J., Salesse R., Combarous Y.;		
RT	"Cloning, sequencing and in vitro functional expression of		
RT	recombinant donkey follicle-stimulating hormone receptor: a new		
RT	insight into the binding specificity of gonadotropin receptors.;"		
RT	J. Mol. Endocrinol. 18:193-202(1997).		
CC	-1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY		
CC	OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE		
CC	ADENYLYL CYCLASE.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
CC	-1- SIMILARITY: TSH SUBFAMILY.		
CC	-1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL: U73659; AAB18245.1; -		

DR HSSP; P23945; 1XUN.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; 7tm_1; 1.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR SMART; SM0013; LRRNT; 1.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 1 687
 FT DOMAIN 18 358
 FT TRANSMEM 359 379
 FT DOMAIN 380 390
 FT TRANSMEM 391 413
 FT DOMAIN 414 435
 FT TRANSMEM 436 457
 FT DOMAIN 458 477
 FT TRANSMEM 478 500
 FT DOMAIN 501 520
 FT TRANSMEM 521 542
 FT DOMAIN 543 565
 FT TRANSMEM 566 589
 FT DOMAIN 590 600
 FT TRANSMEM 601 622
 FT DOMAIN 623 687
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 434 509
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 SQ SEQUENCE 687 AA; 76937 MW; FC3AF0B5531DA9A CRC64;
 Query Match 12.88; Score 492.5; DB 1; Length 687;
 Best local Similarity 25.18; Pred. No. 4.3e-28;
 Matches 167; Conservative 105; Mismatches 274; Indels 119; Gaps 19;
 QY 30 ELGPHNNNIKAIPKAFKMGNPILQTHFYDNPY-QPYGRSAFOYLPKYLHTLSUNGAMDIQ 88
 DB 50 ELRFVLTKLRTVIRKGAFCGFDLKKIEISONDVLEAVFNSLPRLHEIRIEKANN-- 107
 QY 89 EPPDLKGTSLLETLTRAGIRLLPSGMCQQLPRLRVLELSHNOIELPSELHRCCKLEI 148
 DB 108 -----LLYIDHDARQNPINLQYLLISNGIKHLPRAVHKIQSLQKV 147
 QY 149 --GLQHN-RIMEIGADTFQSLS-SIQALDLSWMAIRSHPEAF-----STLH 191
 DB 148 LLDIDQNMIIHYVERNSFGLSESKITRLSKNGIOEHINCAFNQTOLDLNLSDNNNIE 207
 QY 192 SJ-----VKLDLTNOQLTTLPLAGLGLMLKKGKLNALISQARSKDSFKRLTL 240
 DB 208 ELRPNVFOGASGPVILIDISGRISHSLPNYGLNNKKLRARSTYNNLKKPLSLKFAALMBA 267
 QY 241 EPPVAYOCPCPGMCAFPKASQWAEADLHD-----DESSKRPLGLLARAQ 288
 DB 268 SLITYSHCC-----AFANMRQQTSELOTCTCNKSLIRGEVMTQARGERSLAEDD 317
 QY 289 ENHVDQDL-ELQLEMEDSKPHPSVOCSPTPGPKCEYLFESWGIKRLAVAYIVLISVC 347
 DB 318 ESMMSSEFDYDCNEVVD-----VTCSPKPDAPFCEDINGYDILRLVLMFISILATG 371
 QY 348 NGVILLITFAGGAPRLPVPKRVGAIAGANTLTGISGLASVALTFGQFSEYARRET 407
 DB 372 N-IIVLVILITISYKRLTVPRLMCMNLAFADLCIGIYLLIASVDIHTKSYHNVAIDMOT 430

QY 408 GLGCRATGFLAVIGSEASVLLTLAA-----VCCSVSY---SCVRAYGKSPSL 452
 DB 431 GAGCDAAAGFFTVGSELSVYTLTAITLERMHTITHAMOLEKQVLLRRNASMVLGVIJGF 490
 QY 453 GSVRAGVIGLALAGLAALPLASVGEYASPLCLPAPPEGAPALGFTVALVMNNSFC 512
 DB 491 G-----VGL-----LPFGISTYMKVSVICLPN--DIDSPLSOLYMSLILVLA 533
 QY 513 FLVYAGAYIKLYCDUPRGRDF-EAWMDCANVRHVAMLIFADGLLYCPYAFISPMGLFP 571
 DB 534 FVVICGCTHYIYLVTRNPNIVSSSDPKIKAKRMGILITFTPLCMAPISFEGISAKVAL 593
 QY 572 VPEEAVKSVLLVPLPACINPLILFNPHFDDLRRLPRRAGDSGPLAYAAAGELKS 631
 DB 594 IIVSKSKILVLETPYNSCANPFLVAIFPKNFRDFELLSKFG-----CYEMQDTYR 648
 QY 632 SCDSST 636
 DB 649 ETSST 653

RESULT 15
 TSHR_BOVIN
 ID TSHR_BOVIN STANDARD; PRT; 763 AA.
 AC Q27987;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone receptor).
 GN TSHR
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=97280128; Pubmed=9134497;
 RA Silverides D.W., Houde A., Ethier J.F., Lussier J.G.;
 RT "Bovine thyrotropin receptor cDNA is characterized by full-length and truncated transcripts."
 RL J. Mol. Endocrinol. 18:101-112(1997).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC OR send an email to license@isb-sib.ch.
 DB EMBL; U15570; AAC18639.1; -
 DB HSSP; P16473; 1XUM.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 1.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 21
 FT POTENTIAL.

```
FT CHAIN 22 763 THYROTROPIN RECEPTOR.
FT DOMAIN 22 412 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 413 440 1 (POTENTIAL).
FT DOMAIN 441 449 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 450 472 2 (POTENTIAL).
FT DOMAIN 473 493 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 494 516 3 (POTENTIAL).
FT DOMAIN 517 536 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 537 559 4 (POTENTIAL).
FT DOMAIN 560 579 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 580 601 5 (POTENTIAL).
FT DOMAIN 602 624 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 625 648 6 (POTENTIAL).
FT DOMAIN 649 659 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 660 681 7 (POTENTIAL).
FT DOMAIN 682 763 CYTOPLASMIC (POTENTIAL).
FT REPEAT 51 74 LRR 1.
FT REPEAT 125 150 LRR 2.
FT REPEAT 151 174 LRR 3.
FT REPEAT 176 199 LRR 4.
FT REPEAT 201 223 LRR 5.
FT REPEAT 225 248 LRR 6.
FT DISULFID 493 568 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 763 AA; 86431 MW; 35E9F647E7ED7A8C CRC64;
```

Query Match 12.0%; Score 463; DB 1; Length 763;
Best Local Similarity 24.1%; Pred. No. 6.7e-26;

Matches 173; Conservative 118; Mismatches 310; Indels 116; Gaps 19;

```
QY 37 NIKAIPEKAFNGNPLQITHEFYNDPIQFVGSAFOYLPRLHTLSINGAMDIQE----PP 91
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 43 DIQSIPLP-----PSTQTLKFIETHLKITIPSAFSLNPLNISRIYLSIDATIQLESHPY 98
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 92 DLKGTTSLEILTTRAGIRLPSCMCQOLPRLVLELSHNQIEELPSLHR----- 141
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 99 NLSKYTHIEIRN-TRS-LTYIDSGALKELPLKFIPTGTGLRVFPDLTKIYSTDVFTIL 156
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 142 -----CCKLEIETGLOHNRIWEIGADTFESQLSLOALDLSWNA-IRS 181
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 157 EITNPPIWTSIPANAFQGLCNETTLTKLYNNGFTSIQGHAFNG-TKLDAYLNKKYFTV 215
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 182 IHPRAFSTIHS-LVKLDLTDNQTLTPLAGLGGLMHLKLGMLALSOAFSKDSFPKRLIL 240
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 IGQDAFAGVYSGPTLLDISYTSVTALPSKGLEHLKELLARNTWTLPKLPISLFLHTRA 275
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 241 EVPYAYQCCPYG-----MC-----ASFFKASGQWEADDLHLDDEE 275
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 276 DLSYPSHCACARKNOKKIRINGIQLSLMCNESSIRGLRORKSASALNGPFQYEDLIG-DGSA 334
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 276 SSKRPLGLLAQAEENHY-----DQDDEL-----QLEMEDSKRPHPSV----- 312
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 335 GYKENSFKQDTQSNSHYVFEEDDEIIGEQQLKNPOEETLQAFDSHYDYIVCGGSED 394
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 313 -QCSPTPGPFKPCCEYLFPESWIRLAWAIVLLSVLCNGILVLTVEFAGGPAPLPVKFVVG 371
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 395 MVTGPKSDEFNCPEDIMYKFLRIYVWFVSLALLAGNVFVLYLLTSHYKLTVP-RFLMC 453
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 372 AIAGANTLTGISCGLLASVDALTGQFSEYGARWETGLGCRATGFLAVLGSEASVLLTL 431
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 454 NLAFADFCMGYLLLIASVDLYTOSEYYNHAIDMOTGPCNTAGFTYVASELSVYTLTV 513
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 432 AAVGCSVSVGVRAGKPSGLGSAVAGVLCIALAGLAALPLASGEYGAAPLCIPYAP 491
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 514 ITLERMWAITPAMRLDRIRIMHAYVIMLGWVCCFLALLPLVGLISSYAKVSICLPM-- 571
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 492 PEGOPALGFTVALVVMNSFCFLVAVAGAYIKLYCDLPR-----GDFEAVWDCAMVRHVAW 546
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 572 DFTPLALAIYIIVLLIINIAFIYACACVYKITYIVRNPHYNGD-----KDIRIAKRAV 627
```

```
QY 547 LIFADGLICPVAFLESPASMGLFPPVTPBAVKSVLLVYLPACINPLLYLLENPHFRDD 606
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 628 LIFDECMAPISFYALSALMKNKPLITYTNSKILVLFYPLNSCANPFLYLFTRAFORD 687
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 607 LRRIRPAGDSGPLAYAAAGE-----LEKSSCDSTQATVAFSDVDLILEAS 652
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 688 VFMLLSKFGICKRQAQAVRGQRVSPKNSGTGLRVQKVPDPDVQSLPNVQDDYELLENS 744
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 8, 2002, 19:34:59
Job time : 16.7524 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:32:55 ; Search time 18.8167 Seconds
(without alignments)
3760.229 Million cell updates/sec

Title: US-09-851-595-8

Perfect score: 3850

Sequence: 1 GLHNLFTLDLNTYKQLQEPV.....GGLSGGGGFPQSGIAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1844	47.9	907	2	JE0176	orphan G protein-c
2	1759	45.7	907	2	JG0193	G protein-coupled
3	534	13.9	694	2	JC4301	follicle stimulat
4	526.5	13.7	695	2	I45896	follicle stimulat
5	524.5	13.6	695	2	JC1493	follicle stimulat
6	522	13.6	694	2	JC2337	follicle stimulat
7	513.5	13.3	692	2	A34548	follicle stimulat
8	502.5	13.1	695	1	JN0698	follicle stimulat
9	498.5	12.9	695	1	ORHUT	follicle stimulat
10	463	12.0	696	2	JC7361	follicle stimulat
11	460	11.9	764	2	A40077	thyrotropin recept
12	456.5	11.9	696	2	A41344	thyrotropin recept
13	456.5	11.9	764	2	JC5643	thyrotropin recept
14	454.5	11.8	764	2	I48882	thyrotropin recept
15	449.5	11.7	700	2	A42395	thyrotropin recept
16	443.5	11.5	700	2	I77463	thyrotropin recept
17	443.5	11.5	700	2	A49744	thyrotropin recept
18	442.5	11.5	793	3	JC7390	thyrotropin recept
19	441	11.5	764	1	ORHURH	thyrotropin recept
20	432.5	11.2	814	2	JC7389	thyrotropin recept
21	432.5	11.2	889	2	T20213	thyrotropin recept
22	424.5	11.0	764	2	A35956	thyrotropin recept
23	413	10.7	699	1	ORHUT	thyrotropin recept
24	384	10.0	925	1	JC2033	G protein-coupled
25	338	8.8	1115	2	S40241	G protein-coupled
26	248	6.4	1091	2	A58532	glial cell membran
27	240.5	6.2	1066	2	T15864	hypothetical prote
28	238.5	6.2	738	2	T19938	hypothetical prote
29	237	6.2	653	2	T25194	hypothetical prote

30	233.5	6.1	536	2	A34901	lysine carboxypept
31	231.5	6.0	1385	2	T13887	tlr protein - Iti
32	230.5	6.0	1389	2	T13852	gene wheeler prote
33	227.5	5.9	603	2	JC6128	insulin-like growt
34	226	5.9	605	2	JC5239	insulin-like growt
35	226	5.9	605	2	A41915	insulin-like growt
36	224.5	5.8	575	2	T29972	hypothetical prote
37	222.5	5.8	575	2	JC1282	insulin-like growt
38	222	5.8	603	2	AD1822	leucine-rich repea
39	220	5.7	1039	2	T22117	hypothetical prote
40	214	5.6	1531	2	T42218	slit-1 protein hom
41	213	5.5	559	2	T42998	Ras-binding protei
42	206.5	5.4	572	2	T30947	hypothetical prote
43	206.5	5.4	613	2	A88684	protein ACT 2 [imp
44	206	5.4	1523	2	T13953	MeFs protein - ra
45	203.5	5.3	717	2	T33295	hypothetical prote

ALIGNMENTS

Result 1

JE0176 Orphan G protein-coupled receptor precursor - human

C:Species: Homo sapiens (man)

C:Date: 03-Jul-1998 #sequence-revision 10-Jul-1998 #text-change 21-Jul-2000

C:Accession: JE0176

R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A:Title: Identification and cloning of an orphan G protein-coupled receptor of the gl

A:Reference number: JE0176; PMID:98308104; PMID:9642114

A:Accession: JE0176

A:Molecule type: mRNA

A:Residues: 1-907 <MCD>

A:Cross-references: GB:AF062006; NID:g3366801; PIDN:AAC28019.1; PID:g3366802

C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C:Genetics:

A:Gene: H638

F:1-21/Domain: signal sequence

30	233.5	6.1	536	2	A34901	lysine carboxypept
31	231.5	6.0	1385	2	T13887	tlr protein - Iti
32	230.5	6.0	1389	2	T13852	gene wheeler prote
33	227.5	5.9	603	2	JC6128	insulin-like growt
34	226	5.9	605	2	JC5239	insulin-like growt
35	226	5.9	605	2	A41915	insulin-like growt
36	224.5	5.8	575	2	T29972	hypothetical prote
37	222.5	5.8	575	2	JC1282	insulin-like growt
38	222	5.8	603	2	AD1822	leucine-rich repea
39	220	5.7	1039	2	T22117	hypothetical prote
40	214	5.6	1531	2	T42218	slit-1 protein hom
41	213	5.5	559	2	T42998	Ras-binding protei
42	206.5	5.4	572	2	T30947	hypothetical prote
43	206.5	5.4	613	2	A88684	protein ACT 2 [imp
44	206	5.4	1523	2	T13953	MeFs protein - ra
45	203.5	5.3	717	2	T33295	hypothetical prote

Query Match 47.9%; Score 1844; DB 2; Length 907;
Best local similarity 52.2%; Pred. No. 3, le-133;
Matches 360; Conservative 113; Mismatches 186; Indels 30; Gaps 8;

QY	1	GLHNLFTLDLNTYKQLQEPV	IRTLGRLOELGFHNNNKAIPEKAFMGNPLITTHFDN	60
DB	232	GLHSLFTLDLNNINDEPPTAIRLISNKEGFHNNNKIRISPEKAFVGNPSLITTHFDN		291
QY	61	PIQFGRAFOVLPKLTLSUNGAMDIOEPDKGTSELTITRACIRLPSGMCQOL		120
DB	292	PIQFGRAFOVLPKLTLSUNGAMDIOEPDKGTSELTITRACIRLPSGMCQOL		351
QY	121	PLRLVLESHNOIEELPSLHRCOKLEIGLOHNRIMEIGADFFSOLSSQALDLSNNAIR		180
DB	352	PLRLVLESHNOIEELPSLHRCOKLEIGLOHNRIMEIGADFFSOLSSQALDLSNNAIR		411
QY	181	SIHPEAFSLHSLVKLDTNOLITPLAGLGMLHRIKGNLALSQAFSKSPKRLIT		240
DB	412	SIHPEAFSLHSLVKLDTNOLITPLAGLGMLHRIKGNLALSQAFSKSPKRLIT		471
QY	241	EVPYVYQCCPYGMCASFKAQGW-----EADLHLDDESSKRPGLDIAQAENHYDQ		294
DB	472	EVPYVYQCCPYGMCASFKAQGW-----EADLHLDDESSKRPGLDIAQAENHYDQ		520
QY	295	DLDELQLEM-EDSKRHPVQVSCSPFPQPFKPCYELFESWGLRLAWVAIVLISVLCNGVLL		353

```

Db 521 DLEDFLEDEEDLKALHSVOCSPSPGPFKPCHEHLIDGLWLTIGWTVIAVLTALTCALVTS 580
OY 354 TVFAGGPAFLPVKGVYAGTAGANTLTIGISGLLASVDALTFGGFSEYGANWETGLGCR 413
Db 581 TVFR-SPLVSPYSLKLTIGVIAVMMLTGSSAVLAGVADFEGSPARHAGWENGVCHV 639
OY 414 TGFLAVLSEASVLLTLTAVOCSSVSCVAVYKSPSLGSRVAGVGLCLAGLAAALP 473
Db 640 IGFSLTFSESSVFLTLTALBERGFSYKAKFTKAPFSSLKXITILLCALLALMAAVP 699
OY 474 LMSVGEYASPLLPAPPGOPALGFTVALVMNSFCFLVAGAVTKLYCDLPGRDFE 533
Db 700 LAGSKYKASPLCLPL--PGESEPTMGVYVALITLNSLCFLMTIATYTKLYCNDKGDLE 757
OY 534 AVWDCAVRHVAWLIFADGLYCPVAFVAFSPASMLGFPVTPAVKSVLLVLPACLN 593
Db 758 NIMDCSMVKHIALLTFTNCLNCPVAFSLSSSLNTLTFISBEVIKFTLLVVPPLACLN 817
OY 594 LLYLFNPFRRDDLRLRPRA-----GDSGLPAYMAAGELEKSSCDSTQALVAFSDVLI 648
Db 818 LLYLFNPFRRKEDLVSLRKQTYWTRSKHPSLMSINSDDVEKQSCDSTQALVFTSSSIT 877
OY 649 LEASEAGRP-----PGLFTYGPSPVTLISC 673
Db - 878 YDLPSSVSPVAPVATPYESCHLSVAFAVPC 906

```

RESULT 2

```

G:protein-coupled receptor FEK - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JG0193
R:Hermey, G.; Melther, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A:Title: Identification of a novel seven-transmembrane receptor with homology to glycophorin A
A:Reference number: JG0193; MUID:99121227; PMID:9920770
A:Accession: JG0193
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>

```

Query Match 45.7%; Score 1759; DB 2; Length 907;

Best Local Similarity 51.1%; Pred. No. 1e-126; Indels 18; Gaps 8;

```

Matches 349; Conservative 121; Mismatches 195; Indels 18; Gaps 8;
OY 1 GLHLETLIDLVYKNIQEPVPAIRTLGRLQELGFHNNNIKAIFEKAFMGNPLQTHFYDN 60
Db 232 GLHLETLIDLVYNNIDEPPTAIKTLNKEIGFHSNNIRSIERAFVGNPSLITTHFYDN 291
OY 61 PIGVGRSAFOYLKRLHTLSNGAMDIOEPDLKGTSLTSLITLRACIRLLPSGMCQOL 120
Db 292 PIGVGRSAFOHLERLKTITNGASHITFEPHITGTATLESITLGAKISSLPQAVCDL 351
OY 121 PRLVLELSHNOIELPRLRCQRLLEIGLOHNRIMETIGADTFSSLSQALDLSMAIR 180
Db 352 PNLQVLDLSYLNLEEDLSLSCQKIQKIDLRHNEIYEIKGSPQDLFNLRSINLMAWKA 411
OY 181 SHPEARSTHSLVKLTLTNQTLTLPLAGLGGIMHLKLGKALISQAFSKDSPKRLTL 240
Db 412 IIRPNASTPLSLKLDLSSNLSSFPVYGLHGLTHLKLTLNRALOSILPANSFPELKIT 471
OY 241 EYVYAVQCCPYGMCASFKAAGWEADLHDDDESSKRPGLLARQAEHNYDDOLDLEQ 300
Db 472 EMPSAVQCCAFGCGENYKISNMKNKDDGNSVD--DLKKKDGLEQVODE---RLEDFL 526
OY 301 LEM-EDSKPSPVQCCSTPPGPFKCEYLFESWGIRLAVMATVLLSVLNGVLLTVFAGG 359
Db 527 LDEEEDLNALHSVQCSPPGPFKCEHLFGSWLIRIGWTVIAVLTSCNALVALVER-T 585
OY 360 PADLPYKFFVVAIGANTLTIGISGLLASVDALTFGGFSEYGANWETGLGCRATGFLAV 419

```

```

Db 586 PLYISSIKLIGVIAVVDILMGVSAVLAADAFTFEGFAOHGAMWEDGIGCQIVGLSI 645
OY 420 LGEASVLLTLTAAVOCSSVSCVAVYKSPISGVSAGVGLGCLAGLAAALPLASVGE 479
Db 646 PASESSIFLTLTALBERGFSYKSKSFYKAPLPSLAIYLLCVLTATTATPLAGSK 705
OY 480 YGASPLCLPYAPPGOPALGFTVALVMNSFCFLVAGAVTKLYCDLPGRDFEAVMDCA 539
Db 706 YNASPLCLPL--PGESEPTMGVYVALITLNSLCFLMTIATYTKLYCLEKGELENMCD 763
OY 540 MYRHAWLIFADGLYCPVAFVAFSPASMLGFPVTPAVKSVLLVLPACLNPLVLYLF 599
Db 764 MYKHIALTLFANCLYCPVAFVAFSSSLNTLTFISPVIKFTLLVVPPLACLNPLVLY 823
OY 600 NPFRRDDLRLRPRA-----GDSGLPAYMAAGELEKSSCDSTQALVAFSDVLI--LEAS 652
Db 824 NPFRRKEDLVSLRKQTYWTRSKHPSLMSINSDDVEKQSCDSTQALVFTSSSIT 883
OY 653 EAGRP--PGLFTYGPSPVTLISC 673
Db 884 SGASPAYPATESCHLSVAFAVPC 906

```

RESULT 3

```

JG4301
N:Alternate names: follicle-stimulating hormone receptor
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C:Accession: JG4301
R:Remy, J.-J.; Labib-Mansais, Y.; Yerle, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebe
Gene 163, 257-261, 1995
A:Title: The porcine follicle-stimulating hormone receptor: cDNA cloning, functional expression and characterization
A:Reference number: JG4301; MUID:96011644; PMID:7590277
A:Accession: JG4301
A:Molecule type: mRNA
A:Residues: 1-694 <REMA>
A:Cross-references: GB:L31966
A:Experimental source: ovarian granulosa cells
C:Comment: This receptor belongs to the family of the G-protein coupled receptors. It
ermatogenesis in male and oogenesis in female.
C:Genetics:
A:Gene: fshr
A:Map position: 3 q2.2-q2.3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa
C:Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F:366-388/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:398-420/Domain: transmembrane #status predicted <TM1>
F:443-464/Domain: transmembrane #status predicted <TM2>
F:485-507/Domain: transmembrane #status predicted <TM3>
F:528-549/Domain: transmembrane #status predicted <TM4>
F:573-596/Domain: transmembrane #status predicted <TM5>
F:608-629/Domain: transmembrane #status predicted <TM6>

```

Query Match 13.9%; Score 534; DB 2; Length 694;

Best Local Similarity 25.9%; Pred. No. 8.2e-33; Indels 132; Gaps 18;

```

Matches 170; Conservative 106; Mismatches 249; Indels 132; Gaps 18;
OY 30 ELGFHNNNIKAIFEKAFMGNPLQTHFYDNPT-QFVGRSAFOYLKRLHTLSNGAMDIO 88
Db 49 ELRFVYTKRAIDPKKAFSGCDLEKTEISONDVLEAVWFSLPLHEIRIKANLL 108
OY 89 EF-PDLKGTSLTSLITLRACIRLLPSGMCQOLPRLVLELSHNOIELPRLHRCQKLE 147
Db 109 YIDP-----AFQNPRLRYLLISNGVKAHDAVHRIGSLQK 145
OY 148 T--GLQHN-RIMEIGADTFSSLSQALDLSMAIRSHPEARS-----TL 190
Db 146 VLLDIODININITHVERNSFMGLSFESEMIILMSKNGREINCAFNQGDLEMLSDNDL 205
OY 191 HSL-----VKLDITDNLQTLPLAGLGGIMHLKLGKALISQAFSKDSPKRLRI 239

```

Db 206 EELPNDVPHGASGPVILLISRTIRHSLSYGLLENKILKLNARSTYNLKLPLEKEVTLME 265
 Qy 240 LEVPAVYCCCPGCMGASFPKASGMEADLH-----LDE 274
 Db 266 ASLTPSHCCAF-----ANMRQISDLHPICNKSILROEVDMVTOARGORVSLAEQG 317
 Qy 275 ESSKRPGLGLARQENHYDQDLEQLLEMDSKPHPSVOCSPPTGPKCEYLFEWSGIR 334
 Db 318 ESS-----LAKEPDMVSEFNVDLCNEVVD-----VICSPKPDAPNCECDIMGDILR 365
 Qy 335 LAWVAIVLISVLCNGILVLTVPAGCAPLPVPKFEVGAAGANTLTGICGLASDAIT 394
 Db 366 VLIMEFISIAIIGN-IIEVLITTSQYKLTVPFLMCNLAFTDLGILYLLIASIDIFT 424
 Qy 395 FEOFSYGARMEFGICGRATGFLAVLSEASVILLTAA-----VOCSSVY-- 440
 Db 425 KSOYHNHYAINWOTGAGCDAGFFVYFASLSYTTLTATLERKHTTTAMQLOCKVOYIH 484
 Qy 441 -SCVRAYGKSPSLGVSRAVGLCLALAGLAAALPLASVGEYGAAPLCPYPAPPEGPAL 499
 Db 485 AASIMLYG-----WIFAFITVAFPIRGISSYMKVSTCLPM--DIDSPLSQ 527
 Qy 500 GTTVALVMMNSFCFLVAVAGATTKLYCDLPRGD- FEAVWDCAMVRHVAWLIFADGLYCPY 558
 Db 528 LTVSLVLLVNLVAFVVICGYTHIYLVNPNLMSSSDTKIAKRAMLIFTDPLMVPI 587
 Qy 559 AFLPSASMGLPVPPEAVKSVLLVPLPACINPLLYLFNPHFDDLRRLPRAG 615
 Db 588 SFPAISASLKVPLITYSKILLVLFYPINSCANPFLYAITKFRDVFILLSKFG 644

RESULT 4

145896

follicle stimulating hormone receptor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #ext_change 21-Jan-2000

C:Accession: 145896

R.Houde, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.

Mol. Reprod. Dev. 39, 127-135, 1994

A:Title: Structure of the bovine follicle-stimulating hormone receptor complementary DNA

A:Reference number: 145896; MUID:95127199; PMID:7826612

A:Accession: 145896

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-695 <MOD>

A:Cross-references: GB:L22319; NID:g404671; PIDN:AAC37324.1; PID:g404672

C:Genetics:

A:Gene: FSHR

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.7%; Score 526.5; DB 2; Length 695;

Best local similarity 26.1%; Pred. No. 3.1e-32;

Matches 170; Conservative 108; Mismatches 290; Indels 83; Gaps 15;

Qy 30 ELGFHNNNIKAIPKAFMGNPLQTHIFYDNP1-QFYGRSAFOYLPKLTSLNGANDIQ 88
 Db 50 ELRFVILKLVIRKAFSGFDELEKIEISQNDVLEAVFNSLPRKHEIRIKANLL 109
 Qy 89 EF-PDLKGTTSLEILITLRAGIRLLPSGMCQQLPRRLVLESHNOIEELPSLHCQKLE 147
 Db 110 YIDPD-----AFONLPNRLYLLISNGIKHPLAVHKIQSLQK 146
 Qy 148 I--GLOHN-RIMEIGADTFQOLS-SLOALDSWNAIRSIHPEAF-----STL 190
 Db 147 VLIDIDQINIHIVERKSPFGLSPESMTVLSKNGIOEINHCAFNQGLDELINSDSNL 206
 Qy 191 HSL-----VKLIDTNDQLTLLPLAGLGLMLKLGKLNALSQAFSKDSFKLRI 239
 Db 207 EELPNDVPHGASGPVILLISRTIRHSLSYGLLENKILKLNARSTYNLKLPLEKEVTLME 266
 Qy 240 LEVPAVYCCCPGCMGASFPKASGMEADLH-----LDESSKRPGLGLARQ 288

Db 267 ASLTPSHCCAF-----ANMRQISDLHPICNKSILROEVDMVTOARGORVSLAEED 318
 Qy 289 ENHYDQDLEQLLEMDSKPH--SVOCSPPTGPKCEYLFEWSGIRLAWVAIVLISV 346
 Db 319 EPSYAGFVPMSEFNVDLCNEVVDVTCSPBDPAFNPCEDIMGDILRVLIMEFISIAIT 378
 Qy 347 CNGVLLTVFAGCAPLPVPKFEVGAAGANTLTGICGLASVADLTGQFSEYARNE 406
 Db 379 GNILVLY-ILITSQYKLTVPFLMCNLAFTDLGILYLLIASIDVHTKTEHNVAIDMQ 437
 Qy 407 TGIIGCRATGFLAVLSEASVILLTAAVOCSSVSCVRAYGKSPSLGVSRAVGLCLATA 466
 Db 438 TGAGCAGAGFFVYFASLSYTTLTATLERKHTTTAMQLOCKVOYIHVGMIRA 497
 Qy 467 GLAAALPLASVGEYGAAPLCPYPAPPEGPALGFTVALVMMNSFCFLVAVAGATKYCD 526
 Db 498 FAVALLPPIRGISSYMKVSTCLPM--DIDSPLSQYVMSLLVNLVAFVVICGYTHIYLT 555
 Qy 527 LPRGDP-EAVWDCAMVRHVAWLIFADGLYCPYAFLSFASMLLFPVTPRAVSVLLVYL 585
 Db 556 VRPNNTSSSDTKIAKRAMLIFTDPLCMAPISFPAISASLKVPLITYSKILLVLEY 615
 Qy 586 PLPACINPLLYLFNPHFDDLRRLPRAGDSGPLAVYAAGELEKSSCDST 636
 Db 616 PINSCANPFLYAITKFRDVFILLSKFG-----CYEQAQYIRSTST 661

RESULT 5

JC1493

follicle stimulating hormone receptor - sheep

N:Alternate names: follicle stimulating hormone receptor

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 03-Feb-1994 #sequence_revision 15-Feb-1994 #ext_change 21-Jan-2000

C:Accession: JC1493; 147080

R.Khan, H.; Yarney, T.A.; Saitam, M.R.

Biochem. Biophys. Res. Commun. 190, 888-894, 1993

A:Title: Cloning of alternatively spliced mRNA transcripts coding for variants of ovine

A:Reference number: JC1493; MUID:93176195; PMID:8439338

A:Accession: JC1493

A:Molecule type: mRNA

A:Residues: 1-695 <KHA>

A:Experimental source: testis

R.Yarney, T.A.; Saitam, M.R.; Khan, H.; Ravindranath, N.; Payne, S.; Seidah, N.G.

Mol. Cell. Endocrinol. 93, 219-226, 1993

A:Title: Molecular cloning and expression of the ovine testicular follicle stimulating

A:Reference number: 147080; MUID:93351750; PMID:8394255

A:Accession: 147080

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-695 <YAR>

A:Cross-references: GB:L07302; NID:g165884; PIDN:AAA31525.1; PID:g165885

C:Genetics:

A:Gene: FSH-R

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

F;191/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.6%; Score 524.5; DB 2; Length 695;

Best local similarity 26.4%; Pred. No. 4.4e-32;

Matches 175; Conservative 101; Mismatches 273; Indels 113; Gaps 18;

Qy 30 ELGFHNNNIKAIPKAFMGNPLQTHIFYDNP1-QFYGRSAFOYLPKLTSLNGANDIQ 88
 Db 50 ELRFVILKLVIRKAFSGFDELEKIEISQNDVLEAVFNSLPRKHEIRIKANLL 109
 Qy 89 EF-PDLKGTTSLEILITLRAGIRLLPSGMCQQLPRRLVLESHNOIEELPSLHCQKLE 147
 Db 110 YIDPD-----AFONLPNRLYLLISNGIKHPLAVHKIQSLQK 146
 Qy 148 I--GLOHN-RIMEIGADTFQOLS-SLOALDSWNAIRSIHPEAF-----STL 190

Db 147 VLIDIDQININITHVERNSFMGLSFESMIWLSKNGIOEIHNCAFGTQDLDELNSNL 206
 QY 191 HSL-----VKLIDIDNQITLTPLAGLGMHLKLGKGNLALSOAFKSDSPKRLI 239
 Db 207 EELPNDVFOGASGVILDISRTIRLSPLSGLENLKLRASTYLLKLPISLEKVTIVE 266
 QY 240 LEVPYAVOCCPYGCASFASFGMEAEIDL-----LDDESSKRLGLLARQA 288
 Db 267 ASLTPSHCCAF-----ANMRQTSIDLHPICNKSILRQEVDMTQARGRISLAEDD 318
 QY 289 ENHYDQDLDELQLEMEDS--KPHPSVOCSPTPGPKCEYIPEESKIRLAWAIVLVSVL 346
 Db 319 EPSTYAGFDMYSEFDYDLCEVVVDVTCSEPPDAFNPCEDIMGYDILVLIWIFISILAIT 378
 QY 347 CNGLVLTIVAGGPAPLPVKFVVGAGTAGANTLTGISGLASVDALTFEGFSEKAME 406
 Db 379 GNTLVIV-LITSOYKLTIVRFILMCNLAFADLCIGIYLLIASVDVHTKSOYHNVAIDMQ 437
 QY 407 TGLGCRATGFLAVLGSEASVLLTLAAVQ-----CSVSYSC---VRAVYKSPSLGSVRA 457
 Db 438 TGAGCDAAAGFFTFVASELSVYTLTAITLERWHTITHAMQLECKYHVRHAASIMLVGWV-- 495
 QY 458 GVLGCLALAGLAAALPLASVGEYGAAPLCIPYARPEGOPALGFTVALVMNSFCFLVVA 517
 Db 496 -----FAFAVALPFIPTGISSTIKVSIQCLPM--DIDSPISQLYVMSLLVNLVAFVYIC 546
 QY 518 GAYIKLYCDLPGRDF-EAVWDCAMVRHVAMLIFFADGLLYCVAFLSFASMLGFPVTEBEA 576
 Db 547 GCTHYTLVTRNPNITSSSDTKIAKRMAMLIITFDPLCMAPISFPAISASIKVPLITYSK 606
 QY 577 VKSVLVLPPLPCINPLLYLLENPHFRDLRLRPAGDSGLAAYAAGLEKSSCDST 636
 Db 607 SKLLVLEFPYINSCANPFLAIFTKRRDFEFL-----LSKREGYEV 649
 QY 637 QA 638
 Db 650 QA 651

RESULT 6

JC2237

Follicotropin receptor, testis - horse

N:Alternate names: eFSHR

C:Species: Equus caballus (domestic horse)

C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 13-Aug-1999

C:Accession: JC2237; JC2370

R:Robert, P.; Amsellem, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bizard

Biochem. Biophys. Res. Commun. 201, 201-207, 1994

A:Title: Cloning and sequencing of the equine testicular follicotropin receptor.

A:Reference number: JC2237; MUID:94256980; PMID:8198575

A:Accession: JC2237

A:Molecule type: mRNA

A:Residues: 1-694 <ROB>

A:Cross-references: GB:S70150; NID:g546896; PIDN:AA30854.1; PID:g546897

A:Experimental source: testis

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C:Keywords: glycoprotein; hormone receptor; transmembrane protein

F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:366-386/Domain: transmembrane #status predicted <TM1>

F:398-440/Domain: transmembrane #status predicted <TM2>

F:443-464/Domain: transmembrane #status predicted <TM3>

F:485-507/Domain: transmembrane #status predicted <TM4>

F:528-549/Domain: transmembrane #status predicted <TM5>

F:573-596/Domain: transmembrane #status predicted <TM6>

F:608-629/Domain: transmembrane #status predicted <TM7>

F:191,199,268,293/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.6%; Score 522; DB 2; Length 694;
 Best Local Similarity 25.6%; Pred. No. 6,9e-32;
 Matches 168; Conservative 113; Mismatches 282; Indels 92; Gaps 15;

QY 30 ELGFHNNNIKAIEKAFMGNPLQTIHFYDNP1-QEFGRSFQYLPRLHTLSLNGADIQ 88
 Db 50 ELRFVTLKLRVITKGAFFSGGDLKEIISQNDVLEIVANFSLPRLHETRIEKANN-- 107
 QY 89 EPPDLKGTTSLELTTLTRAGIRLLPSGMCQQLPRLVLELSHNOIEBPSLRHCQKLEEI 148
 Db 108 -----LXYIDHDAFONLEPMLQYLLISNIGIKRLPAVHKIQSLQKV 147
 QY 149 --GLOHN-RIMEIGADFEQSL-SIGALDSWNAIRSIHEAF-----STLH 191
 Db 148 LLDIDQININITHVERNSFMGLSFESTTLRSKNGIOEIHNCAFGTQDLDELNSYNLE 207
 QY 192 ST-----VKLIDIDNQITLTPLAGLGMHLKLGKGNLALSOAFKSDSPKRLI 240
 Db 208 EELPNDVFOGASGVILDISGTRHSLPNVGNELMKLRASTYLLKLPISLEKVALMEA 267
 QY 241 EVPYAVOCCPYGCASFASFG-----QMEAEIDLHDESSKRLPG 282
 Db 268 NLTPSHCCAF--ANMRQTSSELQTCNKSILROEYDMTQARGERYSLAEDDESSYPKG 324
 QY 283 LLARQAEHYDQDLDELQLEMEDSKPHPSVOCSPTPGPKCEYIPEESKIRLAWAIVL 342
 Db 325 FDMYSEFYD-----LCNEVD-----VTGSPKPDANPCEDIMGYDILRVLIWIFIS 373
 QY 343 LSVLCLGLVLTIVFAGPAPLPVKFVVGAGTAGANTLTGISGLASVDALTFGEFSEYG 402
 Db 374 LAITGN-IIVLITSOYKLTIVRFILMCNLAFADLCIGIYLLIASVDVHTKSOYHNVA 432
 QY 403 ARWETGICGRATGFLAVLGSEASVLLTLAAVQCSVSYCVRAVYKSPSLGSVAVLGC 462
 Db 433 IDMTGAGCDAAAGFFTFVASELSVYTLTAITLERWHTITHAMQLECKYQVLRHAASVWLVG 492
 QY 463 LALAGLAAALPLASVGEYGAAPLCIPYARPEGOPALGFTVALVMNSFCFLVAGAYIK 522
 Db 493 WIFAFVALLPFIPTGISSTIKVSIQCLPM--DIDSPISQLYVMSLLVNLVAFVYICGYIH 550
 QY 523 LYCDLPGRDF-EAVWDCAMVRHVAMLIFFADGLLYCVAFLSFASMLGFPVTEAVKSVL 561
 Db 551 YLTVRNPNIVSSSDTKIAKRMALITFDPLCMAPISFPAISASIKVPLITYSKILL 610
 QY 582 IYVLPPLPCINPLLYLLENPHFRDLRLRPAGDSGLAAYAAGLEKSSCDST 636
 Db 611 VLFYINSCANPFLAIFTKRRDFILLSKFG-----CYEMQAOQVTRTETIST 660

RESULT 7

A34548

Follicotropin receptor precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999

C:Accession: A34548; A41729

R:Sprengel, R.; Braun, T.; Nikolics, K.; Segaloff, D.L.; Seeburg, P.H.

Mol. Endocrinol. 4, 525-530, 1990

A:Title: The testicular receptor for follicle stimulating hormone: structure and func

A:Reference number: A34548; MUID:91125358; PMID:2126341

A:Accession: A34548

A:Molecule type: mRNA

A:Residues: 1-692 <SPR>

A:Cross-references: GB:J02842; NID:g204183; PIDN:AAA41175.1; PID:g204184

R:Heckert, L.L.; Daley, I.J.; Griswold, M.D.

Mol. Endocrinol. 6, 70-80, 1992

A:Title: Structural organization of the follicle-stimulating hormone receptor gene.

A:Reference number: A41729; MUID:92149579; PMID:1738373

A:Accession: A41729

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-692 <HEC>

A:Cross-references: GB:S81198; NID:g245344; PIDN:AA821415.1; PID:g245345

A>Note: sequence inconsistent with the nucleotide translation

QY 510 SECFLVAGAYIKLYCDLPRGDF-EAVWDCAMVRHVAVLIFADGLLYCPVAFLSFASMLG 50

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

Db 208 ELPNDFHGASGPVILDISRTRHSLPSYGLNLKLRARSTYNLKKLPSEKLVALMEA 20


```

QY 241 EVPRAYOCCEPYGMCASFEPKASGOMEADLH-----LDDESSKRPGLGLARQAE 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 SLATPSHCCAF-----ANMRQISFLHPICKSLIRQGVDMYTORGRSSLAEDNE 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 NHHYDODDELQLEMEDSKRPH--SVQCSPTPGPKPCCEYLFEWSGIRLAWAIVLLSLC 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 SSVSRGFMTEPTEEDYDLCCNEVDYTCSPKPDAPNCPEDIMGYNLRLVILWFISLTLATG 379
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 NGVLVLLVFAGGPAFLPVKPVVGAIAAGANTLGTISGLLSVDALTFGQSEYGARMT 407
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 N-IIVLVLTTSQYKLTVPREFLMCLAPADLCIGIYLLILASVDLHTTSQYINVAIMDQT 438
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 GIGCRATGFLVILGSEASVLLTLTAAYQ-----CSVSVSCVRAVGSPLSGSVRAVIG 461
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 439 GAGCAAGFEFFYFASLSVYTLTATLERMHTITHAMQIDC-----KYOLRHAASVMWG 493
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 462 CLAIAGLAALPLASVGEYASPLCLFPAPRPGOPALGFTVALVMNSFCFLVAVAGAT 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 494 WLFAPFAALPFIPIGSISSYMKVSIPLM--DIDSPISQLYVMSLVLNVLAVVIGCXYI 550
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 522 KLYCDLPKQDF-EAVWDCAMVRHVAMLIFFADGLLYCPVAFSPASMLGLPVYTPPAVKS 580
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 551 HLYLVARNNIYSSSDTYIAKRMAMLIFTDFLCMAPISFFAISASLKLPLIVSKAKIL 610
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 581 LVLVLPPLACLPNLLYLLFNPHFRDDLRLRPRAGSGPLVAAAGELKSSGSDT 636
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 611 LVLFPINSCANPFLVAFITKFRRDFLLSKG-----CYEMQOIYKTETST 661
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
JC7361
folliotropin receptor precursor - newt
N:Alternate names: follicle-stimulating hormone receptor
C:Species: Cynops pyrrhogaster (newt)
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C:Accession: JC7361
R:Nakayama, Y., Yamamoto, T., Oba, Y., Nagahama, Y., Abe, S.
Biochem. Biophys. Res. Commun. 275, 121-128, 2000
A:Title: Molecular cloning, functional characterization, and gene expression of a foll
A:Reference number: JC7361
A:Contents: Testis
A:Accession: JC7361
A:Molecule type: mRNA
A:Residues: 1-696 <NA>
A:Cross-references: DBJ:AB005587
C:Comment: This protein, containing seven transmembrane domains and a large glycosylated
in and thyroid stimulating hormone. This receptor has a common signal transduction pathw
C:Genetics:
A:Gene: fsh-r
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: disulfide bond; glycolysis; glycoprotein; hormone receptor; testis; transmem
F:1-17/Domain: signal sequence #status predicted <SIG>
F:1-17/Domain: follicle-stimulating hormone receptor #status predicted <MAT>
F:18-359/Domain: extracellular #status predicted <TM1>
F:370-389/Domain: transmembrane #status predicted <TM2>
F:402-424/Domain: transmembrane #status predicted <TM3>
F:447-468/Domain: transmembrane #status predicted <TM4>
F:488-511/Domain: transmembrane #status predicted <TM5>
F:523-553/Domain: transmembrane #status predicted <TM6>
F:577-600/Domain: transmembrane #status predicted <TM7>
F:612-653/Domain: transmembrane #status predicted <TM8>
F:616-190,198,267,292/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:445-520/Disulfide bonds: #status predicted

Query Match 12.0%; Score 463; DB 2; Length 696;
Best Local Similarity 24.8%; Pred. No. 2.3e-27;
Matches 158; Conservative 109; Mismatches 283; Indels 86; Gaps 17;

```

```

QY 89 EF-PD-LKGTSTLETLTRAGIRLPLSGMCOQDLRLRLVLEISHNQIHELPSLR--COK 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 YIDPDAFONLPISKYLTLSNIGQLVPA-----VSKIRSFHSLVDV 150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 LEEIGLOHNRIMEIGADTFQSLSLQ-ALDLSMNAIRSHPEAFSTLSIVLDTLQNO- 202
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 QDNINIRH-----IKNSFAGLSSESITIRLNKNIGIEEONHAFNGTH-LNELNLSQNR 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 203 -----LTTLPAGLIGIMHLKLNALQSQAFSKDSEPKLR 238
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 LEKLDPOVQCATGVYIIDISRTIRHFLPNNGCIENIKFRARFNYYLKLPLERFAELI 264
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 ILEVPAYOCCEPYG-----MC-ASFKASGOMEADLH-----DDESSKRP 280
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 EANLTYPHCCAFANERKSEMHPICNKSPFKHSAEPKEDKNLRRESNEDLYSGRS 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 LGLIAGQANHHYDODDELQLEMEDSKRPHSVQCSPTPGPKPCCEYLFEWSGIRLAWAIV 340
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 YSLVNGDEFNNDYILCN--EVHD-----VICPKPDAPNCPEDIMGDNTRLVILWIL 375
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 341 VLLSVLCNGVLLTVFAGGPAFLPVKPVVGAIAAGANTLGTISGILLASVDALTFGQFSE 400
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 SLIATIGN-ITVLVILISQYKFTVPREFLMCLAPADLCIGIYLLILASVDLTKSQIYN 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 401 YGARWETGLCCRTGFLAVLGSSEASVLLTLTAAYQCSVSVSCVRAVGSPLSGSVRAVGL 460
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 HAIDWOTGSGCAAGFEFFYFASLSVYTLTATLERMHTITHAMQIDRKRVRHATIMV 494
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 461 GCLAIAGLAALPLASVGEYASPLCLFPAPRPGOPALGFTVALVMNSFCFLVAVAGAT 520
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 VGMIFAFVVALPFIPIGSISSYTKVSIPLM--DIESRLSOSIYFVLINCAFLILICAC 552
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
A40077
thyrotropin receptor precursor - dog
N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor
C:Species: Canis lupus familiaris (dog)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
C:Accession: A40077; S06933
R:Parmentier, M.; Libert, F.; Maenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; Van S
Science 246, 1620-1622, 1989
A:Title: Molecular cloning of the thyrotropin receptor.
A:Reference number: A40077; MUID:90084524; PMID:2556796
A:Accession: A40077
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-764 <PA>
A:Cross-references: EMBL:X17146; NID:g164098; PIDN:AA330901.1; PID:g164099
A:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <MAT>
F:21-764/Product: thyrotropin receptor #status predicted <MAT>
F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

```

F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:201-225/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match

Best Local Similarity 11.98; Score 460; DB 2; Length 764;
 Matches 167; Conservative 125; Mismatches 331; Indels 100; Gaps 18;

```

17 EPPVAIRTLGRL-----DELGHNNNIKAIEKAMGNPLIQTTHF-YDNPQVGRS 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
36 DEFVTCCKDIRITLPSTPTLTKEITQTLTPSRAPSNLPNISRILYSIDATLQRLSH 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 ARQYLKRLHTLSLNGAMDIQEF-PD-LKGTSSLEILTRFAGIRLPS-GMCOQLRLRY 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
96 SEFNLSKMHIEIRNRSLTSDIPDALKELPLKFLGIFNTGIGVPRDYTKYSTDFEPI 155
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
126 LELSHN-QTEELPS---LHRCCKLEIGLQHNRIWEIGADTSQSLQALDSMNA-IR 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 LELTNPYMASIPANAFQGLCNETTLTKLYNNGFTSIQGHAFNG-TRLDVAVYLNKKYLS 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 SHPEAFSTLHS-LVKLDLTDNQLTTLPLAGLGLMHLKGNLALSOAFSKDSFKLNI 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
215 AIDKDAFGVGGPTLIDVSYTSVTALEPSKGLHKLRLARNTWLTKLPLSLFHLTR 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 LEVPAVYOCOPY-----GMCASF----- 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
275 ADLSTPSHCCAFKKNCKIRGLIESLMCNSSIRSLRQKRSVNTLNGPFDQEEYELGDSH 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
258 --FKASGQM-----EAEDLHLDESSKRPRLGILLARQAEHNYQDLDELQ 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
335 AGYKNSQFOFDNSSHYVFFEEQDEILGQGLKNPOEETLDAPSHDYTV----- 389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
302 EMEOSKHPHSVOCSPTRPGPKCEYLFESWGIRLAVMAIVLTVLNGVLTVPAGGPA 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
390 ----CGNEDWVCTPKSPDEPNCEIDIMGYKFLRYVFWFVLLALLGNVFLVLLTSHYK 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
362 PLPVPKRVVAGIAGANTLTGISCGLASVDALTFGQSEYGARWETGICRATGLAVLG 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
446 LTVP-REFLMCNLAFADPCMGWYLLILIASVDLTHSEYINHAIDMGTGGCTAGFEYFA 504
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
422 SEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSAVAGVGLAGLALAGLAALPLASVGEY 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
505 SELSVYTLTVITLERWVAITFAMRLDRKIRLRHAYAIWVGWVCCFLALLPLVIGISYA 564
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
482 ASPCLLPYAPPEGOPALGFTVALVMNNSFCFLVAVAGAYIKLYCDLPRGDFE-AVWDCAM 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
565 KVSICLPM--DTEPLALATIIIVLNLVIAFTIVCSCYKVIYTVRNPOYNPDCKDTKI 622
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
541 VRIVAMTIFADGLLYCPVAFLSFASMLGLFPVTPPAVKSVLVLPPLACINPLLYLLEN 600
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
623 AKRMAVLIFDFCMGAPISFYALSALMNKPLIVTNSKILLVLFYPLNSCANPLVAFITF 682
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
601 PHERDRLRLRPAGDSGLPAAAG-----ELEKSSCDSTQALVAESVDLIL 649
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
683 KAFQRYVFTLLSKFGICKRQAOYRGQVSPKNSAGIOIKVTRDMROSIPNMQDEYELL 742
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 650 EAS 652
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 743 EMS 745

```

RESULT 12

A11344

Intropin-choriogonadotropin receptor precursor (splice form A) - pig

N:Alternate names: luteinizing hormone-choriogonadotropin receptor

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 21-Jan-2000

R:Accession: A11344

R:Accession: H: Mistrali, M.; Anger, M.; Salesse, R.; Vu Hai-Luu Thi, M.T.; Jolivet, A.;

Science 245, 525-528, 1989

A:Title: Cloning and sequencing of porcine LH-hcg receptor cDNA: variants lacking transmembrane

A:Reference number: A11344; MUID:89332517; PMID:2502844

A:Accession: A41344
 A:Molecule type: mRNA
 A:Residues: 1-696 <100>
 A:Cross-references: GB:M29525; NID:g164528; PIDN:AAA31062.1; PID:g164529
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-696/Product: Intropin-choriogonadotropic hormone receptor #status predicted <MAT>
 F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match

Best Local Similarity 11.98; Score 456.5; DB 2; Length 696;
 Matches 157; Conservative 109; Mismatches 277; Indels 103; Gaps 16;

```

38 IKAIPKAFMG-NPLIQTTHFYNPQVGRSAFYLPKRLHTLSLNGAMDIQEPDLKGT 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 IKYIPGAFGLNEVYKIEISQSDSLKTEANAFDNL-----LNL----- 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
97 TSELTITLRAGIRLPSGMCQQLPRILVLELSHNOIEELPS-----LHRCOK 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
102 ---EILQWTKNLVYIEPGAFTNLPRKYLSTIONTGIRKLPDYTKIFSSSENFILICDN 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145 LE-----EIGQHNRIWEIGADTSQSLQALDSMNA-IRSHPEA 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
159 LHITTPANAFQGMNNESTTLKLYGNGFEEIOSHAFNG-TTILSLEIKENAHLLKWHNDA 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 FSTLSLVKLDLTDNQLTTLPLAGLGLMHLKGNLALSOAFSKDSFKLNILEVPAV 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 FPGAKPSILDISSTYKIALPVGLESITQTLATISYSLKLPSEKFTNLDATLYPS 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 OCCPY-----GMCASFKAAGMEADLHLDESSKRPRLGILLARQAEHNY-----D 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 HCCARNLPTKQNFSTFK-----NFSKQCESTARP-----NNTLYSAIFAE 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
294 QDDELQEMEDSKHPHSVOCSPTRPGPKCEYLFESWGIRLAVMAIVLTVLNGVLL 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 SELSDWDYDGRCP-KTLCQAPEDAFNPCEIDIMGYKFLRYLWNLINILMGWVYLF 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
354 TVPAGGPAFLPVPKRVVAGIAGANTLTGISCGLASVDALTFGQSEYGARWETGICRAT 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
383 VLTSHTKLTVP-REFLMCNLAFADPCMGWYLLILIASVDAQKGYNNHAIDMGTGGCTAG 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
414 TGFLAVLSEASVLLTLTAAVOCSSVSCVRAVYKSPSLGSAVAGVGLAGLALAGLAALP 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
442 AGFTVPAASELSVTLTVITLERWHTITTYALQDLQKRLRHAIPIIMGMLFSTLIMLP 501
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
474 LASVGEYASPLCPYAPPEGOPALGFTVALVMNNSFCFLVAVAGAYIKLYCDLPRGDFE 533
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
502 LVGSSYMKVYSICLPM--DVETTLISQYVILITILNVAVAFIIICACIKIYFANQNEIM 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
534 AV-WDCAMVRHVAMLIRADGLLYCPVAFLSFASMLGLFPVTPPAVKSVLVLPPLACIN 592
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
560 ATNKDKTIAKMAVILFTDTCMAPISEFAISALAKPLITVNSKVLVLFYPVNSCAN 619
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 593 PLIYLFNPHFRDRLRLRPAGDSGLPAAAGELEKSSCDSTQOA 638
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 620 PFLYAITFKAFRRDFLL-----LSKSGCCKHQA 648

```

RESULT 13

JC5643

thyroid stimulating hormone receptor precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 20-Jun-2000

C:Accession: JC5643

R:Bockmann, J.; Winter, C.; Witkowski, W.; Kreutz, M.R.; Boeckers, T.M.

Biochem. Biophys. Res. Commun. 238, 173-178, 1997

A:Title: Cloning and expression of a brain-derived TSH receptor.

A:Reference number: JC5643; MUID:97445147; PMID:9299474

A:Accession: JC5643

A:Molecule type: mRNA

A:Residues: 1-764 <BOC>

A:Cross-references: GB:Y13434; NID:g2462632; PIDN:CAV7846.1; PID:g2462633

A: Experimental source: hypothalamus
 A: Note: the sequences of residues 4-7 and 8-17, 4136-439 and 440-449 are interchanged in C:Comment: This receptor belongs to the guanine nucleotide binding protein (G-protein) c
 C: Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
 F: 1-21/Domain: signal sequence #status predicted <SIG>
 F: 23-764/Product: thyroid stimulating hormone receptor #status predicted <MAT>
 F: 179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
 F: 303-382/Domain: extracellular #status predicted <EXC>
 F: 416-439/Domain: transmembrane #status predicted <TM1>
 F: 451-474/Domain: transmembrane #status predicted <TM2>
 F: 495-518/Domain: transmembrane #status predicted <TM3>
 F: 538-561/Domain: transmembrane #status predicted <TM4>
 F: 581-604/Domain: transmembrane #status predicted <TM5>
 F: 627-650/Domain: transmembrane #status predicted <TM6>
 F: 661-684/Domain: transmembrane #status predicted <TM7>

Query Match 11.9%; Score 456.5; DB 2; Length 764;
 Best Local Similarity 23.9%; Pred. No. 8.4e-27;
 Matches 167; Conservative 116; Mismatches 293; Indels 123; Gaps 20;

37 NIKALPEKAFKGNPLLOTTHFYDNPIDQFGRSAFQYLPKLTLSINGAMDIQF-----PP 91
 43 DIQRPLSP-----PSTQTLKFTETHHKTIPSRASFSLPMSIKYILSIDATIQLESHTY 98
 92 DLKGTSTLEILTRAGIRLPSGMCQQLPRLVLELSHNOIEELPSLRCKLE----- 146
 99 NLSKVTHEIRN-TES-LTYIDSGALKEPLKFLGIFNTGLRVFDPDLTKIYSTDVEFTIL 156
 147 EIGLOHNRWEIGADTFQSLSS-----LQALLDSMNA-IRS 181
 157 EI-TNPNYMTSPVANAFOGLSNETTLTKLYNNGFTSIGHAENGTKLDVYLKKNKYLIV 215
 182 IHPFASTLHS-LVKLDLTDNQLTTLPLAGLGIMHLKGNLALSQAESKDFPKRLIL 240
 216 IQDPAFAGVYSGEPLTDIDISTVTALPNSKGLHKLKELIANRTWTLKPLSLSEHLTRA 275
 241 EWPYAYQCCPY-----GMCASF-KASQW----- 264
 276 DLSYPHSCAFKQKNKIRGLIOSLMCNESIMGLRQKSASALNPGFYOEEDLDGSA 335
 265 -----EADLH-----LDDESS-----KRPGLLAROENHVDODLDELQLE 302
 336 GYKENSFPDTHSNHYVFFEDQDEBITFGQELKNPQETIQAPFNHVDIYVCGSSE 395
 303 MEDSKPHSPVOCSPPTGPFKCEYLFESWGIRLAWAIVLISVLCNGVLVLTFFAGGPAP 362
 396 M-----VCTPKDEFNCPEDINGKFLRIYVWVSLALLGNFVLYILLTSHKL 446
 363 LRPVKEVGAIGANTLTGISCGLASVDALFFGQSEYGARWETIGCRATGFLAVGS 422
 447 TVP-REFLMCNIAFADFCMGVLLILLASVDLTQSEYVNHADIMQTPGCGTAFGFTVFAS 505
 423 EASVLLTLTAOVCSVSGVRAVYKSPSLGSVAGVLCGLALAGLAALPLASVGEYA 482
 506 ELSVYTLVTLEKWAITFAHMLDRKIRLMAHAVIMLGWCCFLALLPLVGISSYAK 565
 483 SPICLPAPPEGQPALGFTVALVMNSFCFLVYAGATIKLYCDLPR-----GDPEAVMD 537
 566 VSTICLPR--DTERPLALAYILLVLLNITAFIYACACVYKITYIVRNPYHNGD-----KD 619
 538 CAMVRHVAWLIFADGLLYCPVALISFASMLGLFPVTPKAVSVLLVPLPACNLPLLYL 597
 620 TRIAKRMAVLLFTDFKMAPISTFALSALMKKPLITVNSKILLVLEPLNSCANPLLYA 679
 598 LFNPHFRDDLRLRPRAGDSGPLAYAAAGLEKSSCDST 636
 680 IFTKAFQDVFMLLSKFGICKRQAQATRGQ-RVSSKNSST 717

RESULT 14
 148882
 thyrotropin receptor precursor - mouse
 N: Alternate names: thyroid-stimulating hormone receptor; TSH receptor

C: Species: Mus musculus (house mouse)
 C: Date: 15-Mar-1996 #sequence revision 15-Mar-1996 #text-change 13-Aug-1999
 C: Accession: I48882
 C: Revision: 148882
 C: Author: S.A.; Oates, E.L.; Hall, C.R.; Grumbles, R.M.; Fernandez, L.M.; Taylor, N.A.
 Mol. Endocrinol. 8, 129-138, 1994
 A: Title: Identification of a point mutation in the thyrotropin receptor of the hyl/hy
 A: Reference number: A54271; MUIID:94224232; PMID:8170469
 A: Accession: I48882
 A: Status: preliminary
 A: Molecule type: mRNA
 A: Residues: 1-764 <RES>
 A: Cross-references: EMBL:U02602; NID:q575923; PIDN:AAB60455.1; PID:q575924
 C: Genetics:
 A: Gene: TSHR
 C: Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
 C: Keywords: G protein-coupled receptor; transmembrane protein
 F: 53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F: 77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F: 102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F: 127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F: 152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F: 179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F: 201-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 11.8%; Score 454.5; DB 2; Length 764;
 Best Local Similarity 22.6%; Pred. No. 1.2e-26;
 Matches 173; Conservative 124; Mismatches 324; Indels 143; Gaps 20;

17 EFPVAIRITGLRL-----QELGFHNNIKAIPEKAFMGNPLQTHF-YDNPIDQFGRS 68
 36 DFRVTCCKELHRIPSLPSTQTLKFTETHHKTIPSRASFSLPMSIKYILSIDATIQLESHP 95
 69 AFQYLPKLTLSINGAMDIQF-PD-LKGTSTLEILTRAGIRLPSGMCQQLPRLR-- 124
 96 SFYNSKMTHEIRNFTSLYIDPDALTELPLKFLGIFNTGLRIFP-----DLTKIYST 150
 125 -----VLELSHN-----QIEELPSLRCKLEIEIGLOHNRWEIGADTFQSLSSQALDSLW 176
 151 DIFILEITDNPYMTSPVANAFOGLCNETTLTKLYNNGFTSIGHAENG-TKLDVAYLNK 209
 177 NA-IRSIHPFASTLHS-LVKLDLTDNQLTTLPLAGLGIMHLKGNLALSQAESKDSF 234
 210 NKYLTAIDNDAFGVYSGPRLTDVSSSTVALPNSKGLHKLKELIATQWTTLKPLSLSF 269
 235 PKLRIEWPYAYQCCPY-----GMCASF----- 257
 270 LHLTRADLSYPSHCCAFKQKNKIRGLIOSLMCNESIRNLRQKSASALNPGFYOEED 329
 258 -----FKASQW-----EADLHLDDESSKRPGLLAROENHVDODL 296
 330 PGDSVGYKQNSKQESPSNHYVFFEDQDEYVGGQELKNPQETIQAFESHYDTV 389
 297 DELQLEMEDSKPHSPVOCSPPTGPFKCEYLFESWGIRLAWAIVLISVLCNGVLVLTFF 356
 390 -----CGDNEDMVCPTKPSDEFNCPEDINGKFLRIYVWVSLALLGNIFVLLILL 440
 357 AGGPAPLPRPKVFGVGAIGANTLTGISCGLASVDALFFGQSEYGARWETIGCRATGF 416
 441 TSHKLTVP-REFLMCNIAFADFCMGVLLILLASVDLTTHSEYVNHADIMQTPGCGTAF 499
 417 LAVGEASVLLTLTAOVCSVSGVRAVYKSPSLGSVAGVLCGLALAGLAALPLAS 476
 500 FTVASSELSTVTLVITLERKWAITFAHMLDRKIRLMAHAVIMLGWCCFLALLPLVWG 559
 477 VGEGASPLCLPYAPPEGQPALGFTVALVMNSFCFLVYAGATIKLYCDL-----PRG 530
 560 ISSYAKVSTICLPM--DIDTPLALAYILLVLLNITAFIYACACVYKITYIVRNPYHNGD 617
 531 DPEANWDCAMVRHVAWLIFADGLLYCPVALISFASMLGLFPVTPKAVSVLLVPLPAC 590
 618 K-----DTRIAKRMVLLFTDFKMAPISTFALSALMKKPLITVNSKILLVLEPLNSC 672
 591 LNPLLYLFNPHFRDDLRLRPRAGDSGPLAYAAAG-----ELEKSSCDSTQAL 639

Db 673 ANPFIATFTAFQDVFITLLSKFGICRQAOAYOGRCVPCNNSTGICQIQKLPQDTRQSL 732
 QY 640 VAFSDVDLILEASEAGRPPELTETGPPSVTLISCOQPGAPRLG 683
 Db 733 PNMDDYELLGNLSQ-----APKIQG 753

RESULT 15

A42395

lutropin receptor - mouse

M/Alternate names: luteinizing hormone-choriogonadotropin receptor

C/Species: Mus musculus (house mouse)

C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999

C/Accession: A42395

R/Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.

J. Biol. Chem. 267, 4479-4488, 1992

A/Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenylyl

more receptor expressed in L cells.

A/Reference number: A42395; MUID:92165799; PMID:1311310

A/Accession: A42395

A/Status: Preliminary

A/Molecule type: mRNA

A/Residues: 1-700 <GUD>

A/Cross references: GB:M81310; GB:M81318; NID:q198811; PIDN:AAA39432.1; PID:q198812

A/Note: Sequence extracted from NCBI backbone (NCBIN:84064, NCBI:84066)

C/Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 11.7%; Score 449.5; DB 2; Length 700;

Best local Similarity 23.1%; Pred. No. 2.6e-26;

Matches 159; Conservative 124; Mismatches 306; Indels 99; Gaps 18;

QY 38 IKAIPKAFMG-NPLQTIHFYDNPQFVGSAPQYLPKLTLSLNGAMDQEPFDLKG 96
 Db 65 VKYIPQAFRGLNEVVKIEISQDSLERIANAFDNL-----LNL----- 105
 QY 97 TSELITLTRAGIRLPSCGCQOLPRLVLELHNOIEELPSLHR-----CQK 144
 Db 106 ---EILIQNTKKNLTYEPGAFTMLPRKYLSICNTGIRTLPDVSKISSSEFFILEICDN 162
 QY 145 L-----EIGIQ--HNRIMEIGADTSPQSLQALDLSMNA-IRSIPEA 186
 Db 163 LYITIPGNAPFGMNNESTITKLYGNGFEVQSHARNG-TTLISLELKENIYTEKMHSGT 221
 QY 187 PSTLSLVKLDLTNOITLPLAGLGLMHLKLGMLALSQAFCSDFFPKRLILEVPYAY 246
 Db 222 FQGATGPSILDVSTRLQALPSHGLESIQTLIATSSYSLKTLPREKFTSLVATLTYP 281
 QY 247 OCCPYGMCASFCKASQWAEADLHDDDESSKRLPLGLARQAEHN--YDQDLDELQ--- 301
 Db 282 HCCAFRNLPK-----KEQNFSESIFFENFSKQCESTVREANNETLYSAIFENELSGW 333
 QY 302 --EMEDSKRHPSPVQSPPTGPKCEYLFEESWIRLAWAIVLSTYLCNGLVLTVPAG 359
 Db 334 DYDYDFCSP-KTLQCTPERDAFNPCEIDINGVAFRLVLMINILATFGN-LTVLFLVLT 391
 QY 360 PAPLPVAFVAGATAGANTLTISGLASVDALTFGQSEYGARWETGLGRATGFLAV 419
 Db 392 RYKLVPRFLKMCNLSFAFCMGLYLLILASVDSQTKGYNHAIDWOTGSGCSAAGFTV 451
 QY 420 LGSEASVLLTFLAAYQCSVSVCAVAYGKSPISGSRAGVGLGCLALAGLAALPLASVGE 479
 Db 452 FASELSVYTLVITLERHNTITTYAVQLODKLRLRAIPIMLGWIPTLMATLPLVGVS 511
 QY 480 YGASPLCLPYAPPEGOPALGFTVALVMNNSFCFLVAVAGATIKLYCDLPRGDEAV-WDC 538

Db 512 YMKVSIQPM--DYESTLSGYIIISILLNNVAFVYICACVRIYFANQNEELTPANKDT 569
 QY 539 AMVHVAMLIFADGLLCPVAFLEFASMLGLEPVYTPKAVKSVLLVLPAPCLNDLTVLL 598
 Db 570 KIAKMAILIFDTQCAPISFFAISAFAKVPDLITVNSKVLTVFYVNSCANPEFLYAV 629
 QY 599 FNPFRDRLRLRPAGDSGFLAVAAAGELKSSODSTQALVAFSDVDLILEASRGRP 658
 Db 630 FTAKFQKDFFLLSRFG-----CCKHRAELYRKREKSACTFNSKNGFPR 673
 QY 659 GLETYGFP-----VYLISCOQPGAPRL 681
 Db 674 SSR-----PSQALILKLSIVHCQOPPPRV 697

Search completed: November 8, 2002, 19:36:53
 Job time : 22.8167 secs

Query Match	28.6%	Score 1103	DB 10	length 631
Best Local Similarity	40.1%	Pred. No. 1.7e-78		
Matches 242	Conservative	96	Mismatches 209	Indels 56
				Gaps 10
QY	LETTTTRAGGATGILLPSGMOQLPRIVLVELSHNOIEELPSLRQKLEELGIGHNRIMEI	158		
	: : : : :			
QY	1 LEMTITLTGKISSINNLCQEGKMLRTLIDLSYNNNRDLPSPNGHALPELSIQNRQIYQI	60		
Db				
QY	159 GADFSQSLSLADLDSNMAIRSHPEAFSTHLSVLRDLTDNOLTTPLAGLGLMHLK	218		
QY	61 KESTFGGLSLRLDLSRDLIHHHSRAATVLGPTTINDVSNFELTSPTEGLNGLMOLK	120		
Db				
QY	219 LKGNLILSQAFSKDSFPEKRLILEVPAYOCQCYGMCASFPAKSGOMEAEEDLHLDFESSK	278		
QY	121 LVGNFKLKELEAKDKDFVNLKLSVPRAYOCCAFWGCDST---AALNEDNSLDQOS---	173		
Db				
QY	279 RPLGLLARQAEHHYDQDLIDLELEEDSKRPHDS--VQCSPTTGPKRPOEYLFESGIRLA	336		
QY	174 -----VADQKFTADAAVNTSTLENEE-HSQIILHCPSTGAFRCFXYLLGSMIMILT	224		
Db				

QY 337 VMAVLTVLVCNGVLTLPVAGCAPLPVYKVVAGTAGANTLTGISCGLASVDALTFPG 396
 Db 225 VWFIFLVALFNNLITITTTTASCTSLPSSKLEIGLISVNLPMGIYTGILFPLDVAWSG 283
 QY 397 QFSEYKARMTGIGCRATGFLAVGSEASVLLTLAAVOCSSVSCVRAVAGSPSLGYSR 456
 Db 284 RFEAFGIMWETGSGCVAGFLAVFSESALFILMLATVERSANDIMKGNHKKOFR 343
 QY 457 AGVLGCLALGALAAALPLASVGEYASPLCLPYAPPEGPAALGFTVALVMNSCEFLV 516
 Db 344 VALLAFIGATVAGCPFLHREYSKSPCLPFP--PTGETPSLGFVTLVLLNSLAFILM 401
 QY 517 AGAYIKLYDLPDRGDEAVWDCAMVRHVAAMLIFFADGLYCPVAFISFASMLGFPVTPDA 576
 Db 402 AVITKLYCMLKEDLSSENSOSSMIKIVAMLIFTNCIFPCVAFSPFAPLITAISSPEI 461
 QY 577 VKSVLLVPLPACLNPLLYLLEFNRD---LRRLPRAGDSGRLAFAAGLEK-- 630
 Db 462 MKSVTLFEPILPACLNPLLYLLEFNRD---LRRLPRAGDSGRLAFAAGLEK-- 630
 QY 631 ---SSCDSTQALVAFSDVD---LILESEAGRPGLIETYPFSPVTLISC 521
 Db 522 YVDCGMVSHLQGNLTVDCCESFLLTKPVSKHLI-----KSHSCPALAVASCOR 569
 QY 674 QQP 676
 Db 570 QRP 572

RESULT 2

US-09-862-767A-4
 ; Sequence 4, Application US/09862767A
 ; Patent No. US20020034786A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCP1C1(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; PRIOR APPLICATION NUMBER: 2001-05-21
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 436
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-862-767A-4

Query Match

Best Local Similarity 20.5%; Score 790; DB 10; Length 436;
 Matches 174; Conservative 73; Mismatches 157; Indels 60; Gaps 9;

QY 240 LEVRYAYOCCPYGMCASFASQWAEADLHDD---EESKRRLGLARQENHYDOD 295
 Db 2 LSPVYAYOCCAFWGCDSY-----ANLNTEDNSLODHSVAGOEKGTADAANTSTPLENDEHSQ 57
 QY 296 LDELQLEMEDSKRHPVQSPPTGPKPCEYLFESWGRILAVMAVTLVLCNGVLTIV 355
 Db 58 I-----IHCPTSTGAFKPKCEYLIGSMIRLIVMFIFFVALFNNLVIIT 103
 QY 356 FAGCAPLPVYKVVAGTAGANTLTGISCGLASVDALTFQFSEYKARMTGIGCRATG 415
 Db 104 FASGTS-LPSSKLEIGLISVNLPMGIYTGILFPLDVAWSGFRFAEFGIMWETGSGCVAG 162
 QY 416 FLAVGSEASVLLTLAAVOCSSVSCVRAVAGSPSLGYSRAGVAGLGLAALPLA 475
 Db 163 FLAVSSSSAIFLMLATVERSANDIMKGNHKKOFRVALLLFLAGTAVAGCPPLF 222
 QY 476 SVGEYASPLCLPYAPPEGPAALGFTVALVMNSCEFLVAGAYIKLYDLPDRGDEAV 535
 Db 401 CESPLTKPVSKHLI-----KSHSCPALAVASCOR 431

Db 223 HRGEYASPLCLPFP--PTGETPSLGFVTLVLLNSLAFILMAVITYTKYCNLEKEDLSEN 280
 QY 536 WDCAMVRHVAAMLIFFADGLYCPVAFISFASMLGFPVTPDAVAVKSVLLVPLPACLNPL 595
 Db 281 SSSMIKIVAMLIFTNCIFPCVAFSPFAPLITAISSPEIMKSVTLFEPILPACLNPL 340
 QY 596 YLLEFNRD---LRRLPRAGDSGRLAFAAGLEK-----SSCDS 635
 Db 341 YVFNPKREKDWKLLKRRVTKSGSVSISSGGCLDODFYDCGMVSHLQGNLTVDCC 400
 QY 636 TQALVAFSDVD---LILESEAGRPGLIETYPFSPVTLISCOOP 676
 Db 401 CESPLTKPVSKHLI-----KSHSCPALAVASCOR 432

RESULT 3

US-09-862-767A-8
 ; Sequence 8, Application US/09862767A
 ; Patent No. US20020034786A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: TANGO-72 AND TANGO-77 NUCLEIC ACID MOLECULES AND POLYPEPTIDES
 ; FILE REFERENCE: MBIO1997-008P1RCP1C1(M)
 ; CURRENT APPLICATION NUMBER: US/09/862,767A
 ; PRIOR APPLICATION NUMBER: 2001-05-21
 ; PRIOR FILING DATE: 1998-08-03
 ; PRIOR APPLICATION NUMBER: US 60/054,646
 ; PRIOR FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-862-767A-8

Query Match

Best Local Similarity 20.3%; Score 783; DB 10; Length 431;
 Matches 173; Conservative 73; Mismatches 157; Indels 60; Gaps 9;

QY 240 LEVRYAYOCCPYGMCASFASQWAEADLHDD---EESKRRLGLARQENHYDOD 295
 Db 2 LSPVYAYOCCAFWGCDSY-----ANLNTEDNSLODHSVAGOEKGTADAANTSTPLENDEHSQ 57
 QY 296 LDELQLEMEDSKRHPVQSPPTGPKPCEYLFESWGRILAVMAVTLVLCNGVLTIV 355
 Db 58 I-----IHCPTSTGAFKPKCEYLIGSMIRLIVMFIFFVALFNNLVIIT 103
 QY 356 FAGCAPLPVYKVVAGTAGANTLTGISCGLASVDALTFQFSEYKARMTGIGCRATG 415
 Db 104 FASGTS-LPSSKLEIGLISVNLPMGIYTGILFPLDVAWSGFRFAEFGIMWETGSGCVAG 162
 QY 416 FLAVGSEASVLLTLAAVOCSSVSCVRAVAGSPSLGYSRAGVAGLGLAALPLA 475
 Db 163 FLAVSSSSAIFLMLATVERSANDIMKGNHKKOFRVALLLFLAGTAVAGCPPLF 222
 QY 476 SVGEYASPLCLPYAPPEGPAALGFTVALVMNSCEFLVAGAYIKLYDLPDRGDEAV 535
 Db 223 HRGEYASPLCLPFP--PTGETPSLGFVTLVLLNSLAFILMAVITYTKYCNLEKEDLSEN 280
 QY 536 WDCAMVRHVAAMLIFFADGLYCPVAFISFASMLGFPVTPDAVAVKSVLLVPLPACLNPL 595
 Db 281 SSSMIKIVAMLIFTNCIFPCVAFSPFAPLITAISSPEIMKSVTLFEPILPACLNPL 340
 QY 596 YLLEFNRD---LRRLPRAGDSGRLAFAAGLEK-----SSCDS 635
 Db 341 YVFNPKREKDWKLLKRRVTKSGSVSISSGGCLDODFYDCGMVSHLQGNLTVDCC 400
 QY 636 TQALVAFSDVD---LILESEAGRPGLIETYPFSPVTLISCOOP 675
 Db 401 CESPLTKPVSKHLI-----KSHSCPALAVASCOR 431

LENGTH: 675
 TYPE: PR
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: deduced sequence
 US-09-877-804-7

Query Match 13.3%; Score 513.5; DB 10; Length 675;
 Best Local Similarity 25.9%; Pred. No. 2e-32;
 Matches 188; Conservative 106; Mismatches 284; Indels 147; Gaps 25;

QY 30 ELGFHNNNIKAIPKAFMGPNLLQTHFYDNP1-QVYGRSAFOYLPKLTLSLNGAMDIO 88
 DB 33 ELRFVILTKLRVLPKGSFAGDLEKIEISONDVLEADVFNSLPKLRHEIRIEKANNL- 91
 QY 89 EFPDCKGTTSLLELTLLTRAGIRLLPSGMCQOLPRLRYLELSHNOIEELPSLHRCQKLEI 148
 DB 92 -----LYINPEAF-QNLPRLRYLLISNTGIKHLPVHKIQLSLQK 130
 QY 149 --GLOHN-RIMEIGADTFESOLS-SLOALDSMAIRSIHPEAF-----STLH 191
 DB 131 LLDIDONINIHIAVARSFMSGLSPEVILWLSKNGIEEIHNCALFNGTQOLDELINLSDNNLE 190
 QY 192 SL-----VKDLTDNQLTTLPLAGIGLMLKLNALSOAFSKDSFPKRLIL 240
 DB 191 ELFPNDVFOGASGVILDISRTKVSHPNGLHNLKLRARSTYRLKLPRLDKFVTLME 250
 QY 241 EYVYAYOCCPYMCASFRRASGOWEADLH-----LDDEE 275
 DB 251 SLTFPSHCCAF-----ANLKRQISELHPICNKSILRODIDDMQIGDQVSLIDEP 302
 QY 276 SSKRPLGLLARQAEHNYDODLDELQLEMEDSKPHPSVOCSPFPKPCFCEYLFEESMGIRL 335
 DB 303 S-----YKGSMDMYNEFDYLCN---EYVD-----VTCSPKPDAPNCEIDMGYILRY 349
 QY 336 AWAIVLLSVLCNGVLLTVFAGGAPRLPVYKVVYGAIAAGANTLTGISCGLASVDALTF 395
 DB 350 LWFISILAITGNTVLYV-VLTTSQYKLTVPFRFLMCNLAADLCIGIYLLIASVDIHTK 408
 QY 396 GGFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAOQ-----CSVSVCVRAYGKS 449
 DB 409 SOYHNVAIDMOTGACDAGGFVFASELVYTLTATTLERHHTITIAMOLEC-----KV 463
 QY 450 PSLSGVRAGVLCGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMN 509
 DB 464 QLRHAASVMVLG-WTFPAFAALPFIIGISSYMKVSICLPM--DIDSPLSQLYVMALLVLN 520
 QY 510 SFCFLVAVAGATYKLYCDLPRGDF-EAVMDCAMVRHVAWLIFADGLLYCPAFLSFASMLG 568
 DB 521 VLAFFVIGCYTHIYLTNRNPTIVSSSDTKIKRMAATLFTDFLCMAPIISFAISASLK 580
 QY 569 LFPVTPKAVSVLLVLLPLACINPLLYLLFNPHFRDRLRLRRAGDSGPLAAAGEL 628
 DB 581 VPLITVSKAKILLVLYFPIYNSCANPFLYALFTKNFRDFIL-----L 623
 QY 629 EKSSCDSTQALVAFSDVDLLEASGAPRPGLETYGFPSVTLLSCQOPAPRLGESHCV 688
 DB 634 SKFGCYEQAOIYRT-----ETSSA-----TNHF-HARKSHCS--SAPRYTNSYLV 667
 QY 689 PEGNH 693
 DB 668 PL-NH 671

RESULT 7
 US-09-877-804-6
 Sequence 6, Application US/09877804
 Patent No. US20020061557A1
 GENERAL INFORMATION:
 APPLICANT: Nikolic, Karoly
 APPLICANT: McFarland, Keith C.
 APPLICANT: Segal, Deborah L.
 APPLICANT: Seeburg, Peter H.

TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
 FILE REFERENCE: P0576P1C2
 CURRENT APPLICATION NUMBER: US/09/877,804
 CURRENT FILING DATE: 2001-09-04
 PRIOR APPLICATION NUMBER: US 08/207,814
 PRIOR FILING DATE: 1994-03-07
 PRIOR APPLICATION NUMBER: US 07/781,153
 PRIOR FILING DATE: 1991-10-31
 PRIOR APPLICATION NUMBER: US 07/347,683
 PRIOR FILING DATE: 1989-05-05
 NUMBER OF SEQ ID NOS: 22
 SEQ ID NO 6
 LENGTH: 692
 TYPE: PR
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: deduced sequence
 US-09-877-804-6

Query Match 13.3%; Score 513.5; DB 10; Length 692;
 Best Local Similarity 25.9%; Pred. No. 2.1e-32;
 Matches 188; Conservative 106; Mismatches 284; Indels 147; Gaps 25;

QY 30 ELGFHNNNIKAIPKAFMGPNLLQTHFYDNP1-QVYGRSAFOYLPKLTLSLNGAMDIO 88
 DB 50 ELRFVILTKLRVLPKGSFAGDLEKIEISONDVLEADVFNSLPKLRHEIRIEKANNL- 108
 QY 89 EFPDCKGTTSLLELTLLTRAGIRLLPSGMCQOLPRLRYLELSHNOIEELPSLHRCQKLEI 148
 DB 109 -----LYINPEAF-QNLPRLRYLLISNTGIKHLPVHKIQLSLQK 147
 QY 149 --GLOHN-RIMEIGADTFESOLS-SLOALDSMAIRSIHPEAF-----STLH 191
 DB 148 LLDIDONINIHIAVARSFMSGLSPEVILWLSKNGIEEIHNCALFNGTQOLDELINLSDNNLE 207
 QY 192 SL-----VKDLTDNQLTTLPLAGIGLMLKLNALSOAFSKDSFPKRLIL 240
 DB 208 ELFPNDVFOGASGVILDISRTKVSHPNGLHNLKLRARSTYRLKLPRLDKFVTLME 267
 QY 241 EYVYAYOCCPYMCASFRRASGOWEADLH-----LDDEE 275
 DB 268 SLTFPSHCCAF-----ANLKRQISELHPICNKSILRODIDDMQIGDQVSLIDEP 319
 QY 276 SSKRPLGLLARQAEHNYDODLDELQLEMEDSKPHPSVOCSPFPKPCFCEYLFEESMGIRL 335
 DB 320 S-----YKGSMDMYNEFDYLCN---EYVD-----VTCSPKPDAPNCEIDMGYILRY 366
 QY 336 AWAIVLLSVLCNGVLLTVFAGGAPRLPVYKVVYGAIAAGANTLTGISCGLASVDALTF 395
 DB 367 LWFISILAITGNTVLYV-VLTTSQYKLTVPFRFLMCNLAADLCIGIYLLIASVDIHTK 425
 QY 396 GGFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAOQ-----CSVSVCVRAYGKS 449
 DB 426 SOYHNVAIDMOTGACDAGGFVFASELVYTLTATTLERHHTITIAMOLEC-----KV 480
 QY 450 PSLSGVRAGVLCGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMN 509
 DB 481 QLRHAASVMVLG-WTFPAFAALPFIIGISSYMKVSICLPM--DIDSPLSQLYVMALLVLN 537
 QY 510 SFCFLVAVAGATYKLYCDLPRGDF-EAVMDCAMVRHVAWLIFADGLLYCPAFLSFASMLG 568
 DB 538 VLAFFVIGCYTHIYLTNRNPTIVSSSDTKIKRMAATLFTDFLCMAPIISFAISASLK 597
 QY 569 LFPVTPKAVSVLLVLLPLACINPLLYLLFNPHFRDRLRLRRAGDSGPLAAAGEL 628
 DB 598 VPLITVSKAKILLVLYFPIYNSCANPFLYALFTKNFRDFIL-----L 640
 QY 629 EKSSCDSTQALVAFSDVDLLEASGAPRPGLETYGFPSVTLLSCQOPAPRLGESHCV 688
 DB 641 SKFGCYEQAOIYRT-----ETSSA-----TNHF-HARKSHCS--SAPRYTNSYLV 684
 QY 689 PEGNH 693

RESULT 9

[illegible]

Db 604 FTKAQRDFLLLSRFG-----CKKRA-----ELYRK 633
 QY 659 GLEFY-----GFP-----SVLLISQGPAPR 680
 Db 633 EFSAYTSCNCKNGFPGASKSQATLKLSTVHQQQIIPR 670

```

RESULT 10
US-09-877-804-2
: Sequence 2, Application US/09877804
: Patent No. US20020061557A1
: GENERAL INFORMATION:
:   APPLICANT: Nikolic, Karoly
:   APPLICANT: McFarland, Keith C.
:   APPLICANT: Segalo, Deborah L.
:   APPLICANT: Seeburg, Peter H.
:   TITLE OF INVENTION: Glycoprotein Hormone Receptor Molecules
:   FILE REFERENCE: P0576P12
:   CURRENT APPLICATION NUMBER: US/09/877,804
:   CURRENT FILING DATE: 2001-09-04
:   PRIOR APPLICATION NUMBER: US 08/207,814
:   PRIOR FILING DATE: 1994-03-07
:   PRIOR APPLICATION NUMBER: US 07/781,153
:   PRIOR FILING DATE: 1991-10-31
:   PRIOR APPLICATION NUMBER: US 07/347,683
:   PRIOR FILING DATE: 1989-05-05
:   NUMBER OF SEQ ID NOS: 22
:   SEQ ID NO 2
:     LENGTH: 700
:     TYPE: PRT
:     ORGANISM: Artificial sequence
:     FEATURE:
:       OTHER INFORMATION: deduced sequence
US-09-877-804-2

```

Query Match	11.5%;	Score 443.5;	DB 10;	Length 700;
Best Local Similarity	22.9%;	Prod. No. 6.3e-27;		
Matches 160;	Conservative 117;	Mismatches 300;	Indels 121;	Gaps 18

Db 512 YKVCISCLP--DVESTLSQVYIIISLILNVAFVVICACIRIFYAQNDELPAKNDF 566

QY 539 AMRVHAWLIFADGLLYCPVAFLEPSASMGLPEVTPKAVKSVLLVWLPACINPELLYTL 598

Db 570 KIAKKMAILITFDETCMAPISEFAISAAFKVPLTIVNSKILLVLPVNSCAPPETIYAI 639

QY 599 FNPHRDRLRLRRAGISGFLATAAAGELKSSCDSTQALVAFSDVDLILEASEGRPP 658

Db 630 FTKAFQDFELLISRFG-----CCKRRA-----ELYRRK 658

QY 659 GLETTY-----GFP-----SVTLISCOOPGAR 680

Db 659 EFSAYTSNCKNGFGASKPSQATLKSLTVCACQDIPPR 696

```

RESULT 11
US-09-928-175-3
; Sequence 3, Application US/09928175
; Patent No. US20020123618A1
;
GENERAL INFORMATION:
;
APPLICANT: Paszly, Christopher J.
APPLICANT: Gong, Jianhua
APPLICANT: Daugherty, Betsy
APPLICANT: Rogers, No. US20020123618A1a
TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
FILE REFERENCE: 00-1229
CURRENT APPLICATION NUMBER: US/09/928,175
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/224,455
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.0
SEQ. ID NO. 3

```

```

;      ; LENGTH: 718
;      ;
;      ; TYPE: PRT
;      ; ORGANISM: Homo sapiens
US-09-928-175-3
```

Db 425 DCLMGVYLFEVGIEDIKRGQYOKYALLMMSVOCRLMGFLAMLSTEVSVLLITLYLTLEK 484

Qy 437 SVSV-----SCVRAKSPSLGSRAGVLCGLALAG-LAAALPLASV-----GEYGASPLCL 487

Db 485 FLVIVFPSPNIRPKGRQTSV-----ILICIMAGFLIAVIPPWNNKDYGFNGKNGVCF 538

Qy 488 P-YAPPEGOPALGFTVALVM-MNSFCPLVAGAYIKLYCDLPRGDPEA--VWDC----- 538

Db 539 PLYDQTEIDIGSKGSLGIFLGVNLAFLIVFSYITMFCISOKTALQTEVANCGRREV 598

Qy 539 AMVRHVAWLIFADGLLYCPVAFLSPASMLGLFPV--TPRAVKS-VLLIVLPLPACINPLLY 596

Db 599 AVANRFFEFIVESDALICWIPFVV---KILSLRVELIPDTMTSWIVIFPLVNSALNPILY 655

Qy 597 LLENHFRDRLRL 610

Db 656 TLTTFNFFKDKLQOL 669

RESULT 12

US-09-928-175-2
 ; Sequence 2, Application US/09928175
 ; Patent No. US20020123618A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher J.
 ; APPLICANT: Gong, Jianhua
 ; APPLICANT: Daugherty, Betsy
 ; APPLICANT: Rogers, No. US20020123618A1ma
 ; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
 ; TITLE OF INVENTION: Uses Thereof
 ; FILE REFERENCE: 00-1229
 ; CURRENT APPLICATION NUMBER: US/09/928,175
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/224,455
 ; PRIOR FILING DATE: 2000-08-10
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 754
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-928-175-2

Query Match 11.1%; Score 425.5; DB 10; Length 754;
 Best Local Similarity 24.2%; Pred. No. 1.8e-25;
 Matches 163; Conservative 105; Mismatches 235; Indels 171; Gaps 24;

Qy 3 HNEETLDLWYKLOEPP--VAIRTLGRLOELGFHNNNIKAIPKAFMGNPILLQTHIFYDN 60

Db 137 NNVTLSTLKKKIHSLPKVYFK-YTKLKKIFLOHNCIRHISRAKAFGLCNLQILYLNHN 195

Qy 61 PLOVGRSAFOYLPRKILHTLSLNGAMDIOEPPDLK-----GTTSEILITLITRAGIRLLP 113

Db 196 CTTTLRPGIFKDLHQTLWILD-----DNPITRISQRLFTGLNSLFFLSVNNVLEALP 249

Qy 114 SGMCQOLPRILVLESHNOIEL--PSLHRCQKLEELGLOHNRITWEIGADTFQSLSLSQA 171

Db 250 KQMCQMPOLNWDDEGNRIKYLNSFTFLSCDSLTVLFLFRNOIGFVPEKTFSSIKNGE 309

Qy 172 LIDSWAITSIHPEAFSTLSLVKLDLTDNQLTTLPLAGLGIMHLKLGKGNLALSQAFSK 231

Db 310 LIDISSTTEIELSHFLKDKLKLQKLNLSNP-----LMTLKNQPEESLKQLOSL 358

Qy 232 DSFPKRLILEVPA-----YQCCPYGMCASFKAQSGQWEADLHLDD 273

Db 359 D-----LERIEIPNITRMFOPMKNLSHIYFKNFRYSYA----- 393

Qy 274 EESKRPGLGLAQAENHYQDDLELOLMEDESKPHPSVQCSPTPGPKPCPEYLEFSMGI 333

Db 394 -----PHVRI-CMPLTLDGSSFDLLANNIL 418

Qy 334 RLAWMAVILLVLCNGVLVLTTFAGGPAPLPVYKFVVGAA--IAGANTLTJISGGLASV 390

Db 419 RIFWVIAFITCFGN-----LEVIGMRSPFKMENTHAMSIKILCA 460

Qy 391 DALT-----FGQFSEYGAHMETGLGCRATGFLAVLQSEASVLLITLAAVOC 436

Db 461 DCLMGVYLFEVGIEDIKRGQYOKYALLMMSVOCRLMGFLAMLSTEVSVLLITLYLTLEK 520

Qy 437 SVSV-----SCVRAKSPSLGSRAGVLCGLALAG-LAAALPLASV-----GEYGASPLCL 487

Db 521 FLVIVFPSPNIRPKGRQTSV-----ILICIMAGFLIAVIPPWNNKDYGFNGKNGVCF 574

Qy 488 P-YAPPEGOPALGFTVALVM-MNSFCPLVAGAYIKLYCDLPRGDPEA--VWDC----- 538

Db 575 PLYDQTEIDIGSKGSLGIFLGVNLAFLIVFSYITMFCISOKTALQTEVANCGRREV 634

Qy 539 AMVRHVAWLIFADGLLYCPVAFLSPASMLGLFPV--TPRAVKS-VLLIVLPLPACINPLLY 596

Db 635 AVANRFFEFIVESDALICWIPFVV---KILSLRVELIPDTMTSWIVIFPLVNSALNPILY 691

Qy 597 LLENHFRDRLRL 610

Db 692 TLTTFNFFKDKLQOL 705

RESULT 13

US-09-928-175-21
 ; Sequence 21, Application US/09928175
 ; Patent No. US20020123618A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher J.
 ; APPLICANT: Gong, Jianhua
 ; APPLICANT: Daugherty, Betsy
 ; APPLICANT: Rogers, No. US20020123618A1ma
 ; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
 ; TITLE OF INVENTION: Uses Thereof
 ; FILE REFERENCE: 00-1229
 ; CURRENT APPLICATION NUMBER: US/09/928,175
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/224,455
 ; PRIOR FILING DATE: 2000-08-10
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 21
 ; LENGTH: 718
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-928-175-21

Query Match 11.0%; Score 423.5; DB 10; Length 718;
 Best Local Similarity 23.7%; Pred. No. 2.4e-25;
 Matches 154; Conservative 109; Mismatches 262; Indels 125; Gaps 19;

Qy 4 NLEETLDLWYKLOEPPVAI-RILGRLOELGFHNNNIKAIPKAFMGNPILLQTHIFYDNPI 62

Db 102 NVTLSTLKKKIHRLPVKVFSTRTELKLYLOHNCITHTSRRAFLGHNQILYLSHNCI 161

Qy 63 QVGRSAFOYLPRKILHTLSL--NGAMDIOEPPDLKGTTSLEILITLITRAGIRLLPSGMCQOL 120

Db 162 TSLRPGTEFQDLQALWLLTDNPTTRISO-KSPWGLNSLFFLSVNGRRLALPETHLCAQM 220

Qy 121 PRLVLELSHNOIEL--PSLHRCQKLEELGLOHNRITWEIGADTFQSLSLSQAIDLSWNA 178

Db 221 POLNWDVLANNGIKYITNSFTFLCDSLTVLFLPRNOIGFVPEKTFSSIKMLGELDLSNM 280

Qy 179 IRSIHPEAFSTLSLVKLDLTDNQLTTLPLAGLGIMHLKLGKGNLALSQA-----FSKDSF 234

Db 281 TKLVPVHLSFDHLQKLNLSNPILYVHRNQGSLKQLO--SLDERIEIPNISTGMF 337

Qy 235 PKRLILEVPA--AYQCCPYGMCASFKAQSGQWEADLHLDESSKRPGLGLAQAENHY 292

Db 338 QPMKNLSHIYLTFRYCSY----- 356

Qy 293 DQDLELOLMEDESKPHPSVQCSPTPGPKPCPEYLEFSMGI RLAWMAVILLVLCNGVLVLT 352

Db 357 -----VPHVRI-CMPSDGISSSEDDLLANGILRSVWVIAFITCVGNFLVI 401
QY 353 LTVFAGGAPLPPVKFYVGAAGANTLTGISGGLASVDALT-----FGQF 398
Db 402 A-----VRSIIKAENTTHAMSIKILCCADCLMGVYLFVSGVDIKYRGQY 446
QY 399 SEXGARETGLGCRATGEFLAVIGSEASVLLITLAAVQCSVS-----SCVRAVKGSPSLGS 454
Db 447 OKVALLMMESVPCRLGFLATLTSTEVSVLLITLLEKFLVIVFPFNSLRCKQQTAVAL 506
QY 455 VRAGVLCGLAAGLAAALPLASV-----GEYGASPLCLPYAPREGQ- AALGFTVALVM-M 508
Db 507 ASLWVGF-----LIAVPTREDYFCGNFYGKNVCPLHYDQADFGSRYSLGIFLV 561
QY 509 NSPCFLVAGAYIKLYCDLPRGDEAV-----WDCAMVRHVAWLIFADGLLYCPVAF 561
Db 562 NLAFVIVISYVTMFCSIHKTALQTAEVRSIHGEVAVANREFVIFSDALCMIPFV 621
QY 562 SFASMLGFLPVTPBAVKS-VLLVVLPLPACINPLLYLFPNFRDRL 610
Db 622 KILSLQV--EIPGTITSMIVFPLPVNSALNPILYTLTTSFFKDKLQ 669

RESULT 14

US-09-928-175-20
; Sequence 20, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928, 175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 20
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-20

Query Match 11.0%; Score 423.5; DB 10; Length 737;
Best Local Similarity 23.7%; Pred. No. 2,5e-25;
Matches 134; Conservative 109; Mismatches 262; Indels 125; Gaps 19;

QY 4 NLETLDTNLYKQLQEPYAI-RTIGRLQELGFHNHNIKAIPKAFMGNDLQTIHFYDNP 62
Db 121 NVTLISLKKKIHRLPKVYSRTYELKTYLQHNCTIHISRAFLGLNLILYISHNCT 180
QY 63 QYGRSAFOYLPKLTLSL--NGAMDIQEPFDLKGTSLELLTLTRAGIRLLPSGMCOL 120
Db 181 TSLRPSIFKDLQHLAMLLDNDPTTRISQ-KSEFGLNSLFLPMVGNRLPALPETLQOM 239
QY 121 PRLVLELSNOIEL--PSLHRCQKLEIGLQHNRTWEIGADFFSOLSSIQALDLSMNA 178
Db 240 POLNWDLANNGIKYITNSTFTLTCDSLTVLPLPNOIGVEPEKTFSSKLNGELDLSNM 299
QY 179 IRSIHPEASTLSLVKLDLTDNQLTLPPLAGLGLMHLKLGMLALSQA-----FSKDSF 234
Db 300 ITKLPLVHLSFDHLQKLNSSNPLLYVHKNOFGSLKOLQ--SLDERIEIPINISGMF 356
QY 235 PKIRILEVPR--AYCCPYGMCASFKAQGWAEADLHDEESSKRPGLGLARQAEENY 292
Db 357 QPKNLSHYLTKTFRCYSY----- 375
QY 293 DQDLDELQLEMEDSKHPVSQCSPTPGPKPCYTLFESMGIRLAVMAIVLLSLVCLNGLYL 352

Db 376 -----VPHVRI-CMPSDGISSSEDDLLANGILRSVWVIAFITCVGNFLVI 420
QY 353 LTVFAGGAPLPPVKFYVGAAGANTLTGISGGLASVDALT-----FGQF 398
Db 421 A-----VRSIIKAENTTHAMSIKILCCADCLMGVYLFVSGVDIKYRGQY 465
QY 399 SEXGARETGLGCRATGEFLAVIGSEASVLLITLAAVQCSVS-----SCVRAVKGSPSLGS 454
Db 466 OKVALLMMESVPCRLGFLATLTSTEVSVLLITLLEKFLVIVFPFNSLRCKQQTAVAL 525
QY 455 VRAGVLCGLAAGLAAALPLASV-----GEYGASPLCLPYAPREGQ- AALGFTVALVM-M 508
Db 526 ASLWVGF-----LIAVPTREDYFCGNFYGKNVCPLHYDQADFGSRYSLGIFLV 580
QY 509 NSPCFLVAGAYIKLYCDLPRGDEAV-----WDCAMVRHVAWLIFADGLLYCPVAF 561
Db 581 NLAFVIVISYVTMFCSIHKTALQTAEVRSIHGEVAVANREFVIFSDALCMIPFV 640
QY 562 SFASMLGFLPVTPBAVKS-VLLVVLPLPACINPLLYLFPNFRDRL 610
Db 641 KILSLQV--EIPGTITSMIVFPLPVNSALNPILYTLTTSFFKDKLQ 688

RESULT 15

US-09-804-626-6
; Sequence 6, Application US/09804626
; Patent No. US20020128190A1
; GENERAL INFORMATION:
; APPLICANT: Lobel, Leslie
; APPLICANT: Lustbader, Joyce
; TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN
; FILE REFERENCE: 0575/62259/JPM/SHS
; CURRENT APPLICATION NUMBER: US/09/804, 626
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 6
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-804-626-6

Query Match 10.9%; Score 420; DB 10; Length 699;
Best Local Similarity 24.0%; Pred. No. 4.3e-25;
Matches 148; Conservative 109; Mismatches 286; Indels 74; Gaps 15;

QY 38 IKAIPKAFMG-NPLIQTIHFYDNPQYGRSAFOYLPKLTLSLNGAMDIOEPFDLKT 96
Db 61 VVIPSQAFRGINEVYIKELISQDSLSERIEANAFDNL-----LNLS----- 101
QY 97 TLELITLTRAGIRLLPSGMCOLPRLVLELSHNOIELPS-----LHRCQK 144
Db 102 ---ELLIQNTKMLRYTEPGAFLNLPRLKYLSTICNTGIRFPDVTYKVSSENFLEICDN 158
QY 145 IE-----EIGLQHNRTWEIGADFFSOLSSIQALDLSMNA-IRSIHPEA 186
Db 159 LHITLTPGAFQGMNNESTYIKYNGFEEYOSHANG--TTLTSLLEKEVHLEKHNGA 217
QY 187 FSTLSLVKLDLTDNQLTLPPLAGLGLMHLKLGMLALSQAQSFKSPKIRILEVPR 246
Db 218 FRGAGPKTLDISSSRKQLALPSYGLESIORLATSSYSKLKLPSSREFVNLBATLTYP 277
QY 247 OCCPYGMCASFKAQGWAEADLHDEESSKRPGLGLARQAEENH--YQDDDELQLEME 304
Db 278 HCCA-----FRNLPTKQNFSHSISENFSKO-CESIVRKVNNKTLYSMAESELG 329
QY 305 DSK-----PHPSVQCSPTPGPKPCYTLFESMGIRLAVMAIVLLSLVCLNGLYL 359
Db 330 DYEGFCLP-KTPRCAPREDANPCEDINGYDFLRLVINTLALMGN-MYVLEFLVTS 387
QY 360 PAPLPPVKFVGAIGANTLTGISGGLASVDALTFGQFSEXGARETGLGCRATGFLAY 419

Db 388 RYKLTVPREFLMCNLSFADFCMGLYLLLIASVDSQTKGQYYNHAIDWQTSGCCSTAGFTY 447
QY 420 LGSEASVLLLTIAAVOCSSVSVSVRAYGKSPSLGSRAGVLGCLALAGLAALPLASVGE 479
448 FASELSVYTLVITLERMHTTYAITHDQKRLRHAIIIMLGWLFSSLIAMLPVGVSN 507
QY 480 YGASPLCLPYAPPEGQPALGFTVALVMNNSFCFLVYAGATIKLYCDLPRGDEAY-WDC 538
Db 508 YMKVSIQPPM--DVETTLISQYIILTLILNVVAFELICACYIKIYFAVRNRPMLMATNKDT 565
QY 539 AMVRHVAMLIIPADGILYCPVAFLSFASMLGFPVTPPEAVKSVLLVPLPACLNPLLYLL 598
Db 566 KIAKMALLITFTDFTCMAPISFPAISAAKVPPLITYTNSKVLIVLFYPIINSCANPFLYAI 625
QY 599 FNPHERDDLRLRPARG 615
Db 626 FTKTFQRDFLLSKFG 642

Search completed: November 8, 2002, 19:37:49
Job time : 12.6019 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 19:33:15 ; Search time 15.0533 Seconds
(without alignments)
1438.571 Million cell updates/sec

Title: US-09-851-595-8

Perfect score: 3850
Sequence: 1 GLHNLETLIDLNKYLQEFV.....GGLSGGGGFGPSGLAFASHV 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued_Patents-AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/BACKFILES1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
- and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	29.6	644	2	US-08-866-757-2
2	1140	29.6	644	4	US-09-153-593-2
3	513.5	13.3	632	4	US-07-757-342D-6
4	498.5	12.9	695	1	US-08-487-886-2
5	498.5	12.9	695	3	US-08-482-855-2
6	498.5	12.9	695	4	US-08-474-986-2
7	468	12.2	764	4	US-07-741-453A-60
8	460	11.9	764	4	US-07-741-453A-54
9	457.5	11.9	696	4	US-07-757-342D-4
10	445	11.6	764	4	US-07-741-453A-61
11	444	11.5	764	4	US-07-757-342D-5
12	443.5	11.5	700	4	US-07-757-342D-3
13	441	11.5	764	4	US-07-741-453A-59
14	413	10.7	696	4	US-07-757-342D-10
15	413	10.7	699	4	US-07-757-342D-2
16	392.5	10.2	792	4	US-07-741-453A-56
17	374	9.7	611	4	US-07-757-342D-8
18	374	9.7	636	4	US-07-757-342D-7
19	373	9.7	764	4	US-07-741-453A-29
20	364	9.5	795	4	US-07-741-453A-55
21	318.5	8.3	420	4	US-08-795-876-33
22	318.5	8.3	423	4	US-08-795-876-38
23	318.5	8.3	436	4	US-08-795-876-2
24	313	8.1	336	1	US-08-118-270-54
25	313	8.1	336	5	PCT-US93-08528-54
26	289.5	7.5	327	1	US-08-118-270-55
27	289.5	7.5	327	5	PCT-US93-08528-55

28	250.5	6.5	332	1	US-08-118-270-53	Sequence 53, Appl
29	250.5	6.5	332	5	PCT-US93-08528-53	Sequence 53, Appl
30	248	6.4	1091	3	US-08-986-485-5	Sequence 5, Appl
31	241.5	6.3	1101	3	US-08-986-485-2	Sequence 2, Appl
32	226	5.9	605	1	US-08-190-802A-49	Sequence 49, Appl
33	226	5.9	605	4	US-09-063-950-5	Sequence 5, Appl
34	226	5.9	605	4	US-08-477-346-49	Sequence 49, Appl
35	226	5.9	605	4	US-08-473-089-49	Sequence 49, Appl
36	226	5.9	605	4	US-08-487-072A-49	Sequence 49, Appl
37	222.5	5.8	603	1	US-08-190-802A-50	Sequence 50, Appl
38	222.5	5.8	603	4	US-08-473-089-50	Sequence 50, Appl
39	222.5	5.8	603	4	US-08-487-072A-50	Sequence 50, Appl
40	222.5	5.8	603	4	US-09-191-647-2	Sequence 2, Appl
41	218	5.7	1523	3	US-09-540-245A-2	Sequence 2, Appl
42	218	5.7	1525	4	US-09-540-153-2	Sequence 2, Appl
43	218	5.7	1525	3	US-08-986-485-6	Sequence 6, Appl
44	207	5.4	353	3	US-08-986-485-6	Sequence 6, Appl
45	205.5	5.3	673	4	US-09-063-950-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-866-757-2
Sequence 2, Application US/08866757
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOUBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSM, DEREK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CA71)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAYNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23, 031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-757-2

Query Match 29.6%; Score 1140; DB 2; Length 644;
Best Local Similarity 40.0%; Pred. No. 7.3e-94;
Matches 247; Conservative 96; Mismatches 214; Indels 60; Gaps 9;


```
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
FEATURE: hydrophobic, about 20-23 amino acids in length
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2
```

Query Match Best Local Similarity 12.9%; Score 498.5; DB 1; Length 695;
Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

```
QY 30 ELGFNNNNIAIEKAFMNPLOTHTFYDNP-QEYGRSAFOYLPKHLTLSENGADIQ 88
DB 50 ELRFVLTAKLVKIOKAFSGFGDEKLEISONDVLEADVFNLSRIHRIKRNANL- 108
QY 89 EFPDLKGTTSLETLTLTRAGIRLLPSGMCQQLPRLVLESHNOIELSLHCOQLEI 148
DB 109 -----LYINPEAF-QNLPNLOYLLISNGIKHLDVHRIHSLQKV 147
QY 149 --GLOHN-RIMEIGADTSQLS-SLOALDSNNAIRSHPEAF-----STLH 191
DB 148 LLDIDONINHIETKNSVGSFESVITMLNKGIOEIHNCAPNGQDELMLSDNNNIE 207
QY 192 SL-----VKLDLTNOQLTLPPLAGLGLMHLKLGMLALSOAFSKDSFKRLTL 240
DB 208 ELPPDVPHGASGPVLLDISRTIRHSLPSYGLLENKLLARSTYNNLKLPLEKVALMEA 267
QY 241 EPPAYOCCPYGMCASFKAAGQWEADLH-----LDDESSKRLPLGLLARQAF 289
DB 268 SLVPSHSCAF-----AMRRQISELHPICKNSILRQVDYMTQRRGSSLAENE 319
QY 290 NHDODIDLEQLEKEDSKPMP--SVQSPTPGPKPEYLFESGIRLAWAIVLVLVC 347
DB 320 SSVSRGDMTYTEFDYDLCEVVDVYTCSPKDAFNPCEIDINGVILVLMFISILATIG 379
QY 348 NGVLTLTFVGGAPLPPEVVGATAGANTLIGISGGLASVDALTFGQFSEGARRET 407
DB 380 N-LIVLVLTSSQYKLVPRFLKCNLAFAFDLCIGIYLLIASVDIHHTKSQYHNVADIMOT 438
QY 408 GIGCRATGFLAVGSASVLLTLAAVQ-----CSVSCVNAVAKSKPSLGSVRAVGVLG 461
DB 439 GAGCDAAGFFVAFSELASYTLTAITLERMHTTHAMQDC-----KVQLRHAASVVMVG 493
QY 462 CLALGALAAALPLASGEYASPLCLPYAPRPGQALGFTVALVMMNSFCELVAVGYI 521
DB 494 WL-FAPAAALFPIGISTYMKVSIQPM--DIDPSLSQLYVMSLVLAVLAVVVICGCI 550
```

```
QY 522 KLYCDLPBGDF-EAYWDCAMVHNAVLFAQGLYCPVAFLSFASMLQFLFVTPFAVSV 580
DB 551 LHVLRNPNIVSSSDRIKRNAMLFTDFELCMAPISFAISASLKVPLITYSKAIL 610
QY 581 LTVVLPACNPLLYLLENPFRRDULRRLPRAGDSGLAYAAAGELKSSCST 636
DB 611 LVLPHPINSCANPLLYAIFTKNFRDRDFILLKSG-----CYEMQAOIYRTETIST 661

RESULT 5
US-08-482-855-2
; Sequence 2, Application US/08402855
; Patent No. 6121016
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweickhardt, Rene Lynn
; APPLICANT: Cheng, Shirley Yui Yen
; APPLICANT: Nugent, No. 6121016een Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan P. Williams,
; ADDRESSEE: Ares-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,855
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: 15-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
; TELEFAX: (617) 723-8923
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 695
; TYPE: Amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: -17 to -1
; IDENTIFICATION METHOD: hydrophobic
; FEATURE:
; NAME/KEY: putative amino-terminal extracellular domain
; LOCATION: 1 to 349
; IDENTIFICATION METHOD: similarity with other
; IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
; IDENTIFICATION METHOD: domains, hydrophilic
; FEATURE:
; NAME/KEY: transmembrane domain
; LOCATION: 350 to 613
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
; FEATURE:
; NAME/KEY: putative transmembrane region I
; LOCATION: 350 to 370
; IDENTIFICATION METHOD: similarity to other G
```

```

IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-482-855-2

Query Match      12.9%; Score 498.5; DB 3; Length 695;
Best Local Similarity 24.8%; Pred. No. 4.1e-36;
Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

QY 30 ELGFHNNNIKAIPKAFMGNPLQITIFNDPI-QFVGRSAFOYLPKLTLSLNGAMDIO 88
DB 50 ELRFVLTKLRYOKGAFSGFDLEKIEISQNDVLEADVFSNLPKLTREIRKANL- 108
QY 89 EFPDLKGTSLLEILLTRAGIRLLPSGMCQQLPRLRYLELSHNOIEELPSLRHOCLEEI 148
DB 109 -----LYINPEAF-QNLPNIOYLISNTGKHLPDVHKHISLQKV 147
QY 149 --GLOHN-RIMEIGADTFQSLS-SLQALDISMAIRSHPEAF-----SLH 191
DB 148 LLDIQDININIHTEIRNSFVLSFESVILMLNKNGIOETHNCAFNSTQDELNLSDNNNIE 207
QY 192 SL-----VKLDLTDNOLITPLLAGLGLMLKLGKLTALSOAFSKDSFPRIL 240
DB 208 ELPLNDVHGASGPVLLDSTRIRHSLPSYGLLENLKLKLRARSYNNLKLPTLEKLVAMEA 267
QY 241 EVPYAYOCCPYGMCASFPGASQGWAEEDLH-----LDDESSKRPGLGLARQAE 289
DB 268 SLTYVSHCAAF-----ANMRQISELHPICNKSILRQEDVYMTQFGRSLSAEDNE 319
QY 290 NHYDDDLDELQEMDSRPH--SVQCSPTGPRPKPCETLFFSGWGLRVAWMAIVLTVLC 347
DB 320 SSSYSGFDMTYTEFYDLCNEVVDVTCSPKPAFNPCEDEIMKYNILRLVIMTISLAIIG 379
QY 348 NGVLVLTYPAGRPALPVPKFEVGAAGANTLTGISGLLASVDALTEGOFSEYGARMET 407

```

```

DB 380 N-IIVLITTSQYKLTVPRLMCMIAFADLCIGYLLILASVDIHTFSQTHNYAIDMQ 438
QY 408 GIGCATGFLAVIGSEASVILLETAAVQ-----CSVSCYRAVCKSPSLGVSAGVIG 461
DB 439 GAGCAAGFTYFASLSLYITLITLEKHTITAMOLD-----KVQLRAASVYMG 493
QY 462 CLALGLAALPLASVGEASPLCPYAPREGOPALGFYALVAMNSFCGLYVAGNI 521
DB 494 WI-FAPAAALPFIIGISSYMKVSLCPM--DIDSPLSQLYVMSLLVNLAVVLCGCI 550
QY 522 KLYCDLPRDF-EAWDCAMVHVMILFEADGLICPVAFSLFSPASMLGLFPYTPAVKSV 580
DB 551 HYLTVRNPNIVSSSDTRIARKMAMLTFTDLCNAPISFPALSALXVPLTVSKATL 610
QY 581 LLYVLEPLPACLPPLLYLFPNPHFDRLRLRPAGDSGLPAAVAAAGELKSSCDST 636
DB 611 LVLFHINCANPFLAIFTKNRDFELLSKCG-----CYEMQAIKVTERTST 661

RESULT 6
US-08-474-986-2
GENERAL INFORMATION:
APPLICANT: Kellon, Christie Ann
Schweickhardt, Rene Lynn
Cheng, Shirley Wui Yen
Nugent, No. 6372711een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massll via kermil to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,986
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-8923
TELEFAX: (617) 723-8923
LOCATION: 614 to 678
SEQUENCE DESCRIPTION: SEQ ID NO: 2;
US-08-474-986-2

Query Match      12.9%; Score 498.5; DB 4; Length 695;
Best Local Similarity 24.8%; Pred. No. 4.1e-36;
Matches 163; Conservative 116; Mismatches 284; Indels 93; Gaps 18;

QY 30 ELGFHNNNIKAIPKAFMGNPLQITIFNDPI-QFVGRSAFOYLPKLTLSLNGAMDIO 88
DB 50 ELRFVLTKLRYOKGAFSGFDLEKIEISQNDVLEADVFSNLPKLTREIRKANL- 108
QY 89 EFPDLKGTSLLEILLTRAGIRLLPSGMCQQLPRLRYLELSHNOIEELPSLRHOCLEEI 148
DB 109 -----LYINPEAF-QNLPNIOYLISNTGKHLPDVHKHISLQKV 147

```


Sat Nov 9 13:04:31 2002

us-09-851-595-8.rai

page 7

APPLICANT: DIMONT, JACQUES
APPLICANT: VASANT, GILBERT
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
TITLE OF INVENTION: 62
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARDY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,453A
FILING DATE: 19911015
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-741-453A-54

Query Match 11.9%; Score 460; DB 4; Length 764;
Best Local Similarity 23.1%; Pred. No. 1.4e-32; Indels 100; Gaps 18;
Matches 167; Conservative 125; Mismatches 331;

QY 17 EPPVARTLGR-----QELGFHNNKAIPEKAFMGNPLLOTTHF-YDNPIDFVGS 95
DB 36 DFRVYCKDHRITPLPPSTQTLKFIETQKTPSRASNLPNISRYISIDATLQRESH 125
QY 69 AFQYLPKLTLSINGAMDIOEF-PD-LKGTTSLETLITRAGITLLPS-GMCOOLPRLRV 155
DB 96 SFYNLSKMTHTIEIRNTRSLSTIDPALKELPLKELGIFNGLGVFPTVTVYGVDFEFI 180
QY 126 LELSHN-QIEELPS--LHRCOKLEELGLOHNRIMEIGADPFSQSLSDALDLSMNA-IR 214
DB 156 LEIDNPMASISIPANAFQGLCNETLTKLYNNGFTSIQHAENG--TKDAVILNKKKLS 239
QY 181 SIHPAFSTLS-LVKLDLTNDNLTPLAGLGLMLHKKKGLALSOAFSKDSFPKRLI 274
DB 215 AIDKDFEGVYSGPTLDVSTYVALPSKGLHKLIRANWTLLKKLPISLSFHLNR 257
QY 240 LEVPAVYCCPY-----GMCAFP----- 334
DB 275 ADLSYPSHCACAFKNOKIRIGLIESLMCNESIRSLRQKSVNLTNGFPDDEYELGDSH 301
QY 258 --FKASQW-----EADLHLDDESSKRPGLLARQAEHNYDQDLDELQ 389
DB 335 AGYADNQGQDDTNSHYVFEEDDELIGQGLKLNQOEETLOAFSDSHDYTV----- 389
QY 302 EMEDSKHPVQSGPMPKPEYLFESMGIRLAVAILVSLCNGVLLTVFAGGPA 361
DB 390 ---CGMEDVCTPKSDENPCEDINGYKFLRIYVWFVSLALGNVVLVLTSHYK 445
QY 362 PLPVPKVVGAAGANTLTIGISGLASVDALFFGQFSEYGARWETGLCCRATGELAVIG 421
DB 446 LVVP-RFLMCNLAFADFCMGMYLILLASVDLTHSEYNNHAIMDQMGPGCMTAGFTTVA 504

QY 422 SEASVLTFLAAVOCVSVSCVRAYKSPSGSVRAGVGLCTALAGLAADPLASVBYG 481
DB 505 SELSYTTLVITLERVAITFAMRLDRIRRAHVAIMGWCCFLALPLVGISSIA 564
QY 482 ASPLCIPIAPBPQGAALGFTVALVMNSFCFLVACATIKYLCDDPRGDFE-AWDCAM 540
DB 565 KVSICLPM--DETPLALXILLVLLNVAIFIVCSGVYKIIYVANNPPYNGDDYFI 622
QY 541 YRHVAMLIADGLLYCPVAFISFASMLGIFVTPKAVSVLLVPLPACLNPLVLEFN 600
DB 623 AKRAVLIETDFWCMAPISFYALSALNKPILITVNSKILLVLFYPLNSCANPLVAFIT 682
QY 601 PHFPDRLRLRPAGDSGPIAVAAAG-----ELEKSCDSTQALVAESVDLIL 649
DB 683 KAFQDVFILLSKGICKRQAOAYRQORVSPKNSAGTQKQVTRMROSLPMDDEVELL 742
QY 650 EAS 652
DB 743 EMS 745

RESULT 9
US-07-757-342D-4
Sequence 4, Application US/07757342D
Patent No. 6218509
GENERAL INFORMATION:
APPLICANT: IGARASHI, Masao
MINAGISHI, Takashi
NAKAMURA, Kazuo
TITLE OF INVENTION: NAKAMURA, PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/757,342D
APPLICATION NUMBER: 31003
FILING DATE: 10-Sep-1991
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BOCKLEY, Linda M.
REGISTRATION NUMBER: 41226
REFERENCE/DOCKET NUMBER: 41226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-07-757-342D-4

Query Match 11.9%; Score 457.5; DB 4; Length 696;
Best Local Similarity 24.2%; Pred. No. 2e-32; Indels 103; Gaps 16;
Matches 159; Conservative 109; Mismatches 285;

QY 28 LOELGFHNNKAIPEKAFMG-NPLQTFHFYDNPIDFVGRSAFOYLPKLTLSINGAMD 101
DB 51 LSRSLITLITVIVISQAEKRGINEVYKTEISQSDSLEKTEANADNL-----LNTS----- 101


```

QY 87 IOEPDLKGTSTLFLTLTRAGIRLPSGMCODLRLVLESHNOIEPS----- 138
Db 102 -----EILIONKRNLYIEPGAFNLPRLKLTICNTGIRLPPVTRFISSE 148
QY 139 -----LHRCOKLE-----EIGQHNRIWEIGADPESQSLQALDLSMN 177
Db 149 FNFLEICDHLITTPANAFQOMNESTTLKYGCEEEIOSHAFNG-TTLISLEKEN 207
QY 178 A-IRSHPEAFSTLSLVLDLTDNLTTLPLAGLGIMHKLKGNLNAQSKDSFEPK 236
Db 208 AHLKKNHNDAFRGARPSILDISSITKQLAPSYGLESITLIANSYSLKLPSEKFTN 267
QY 237 LRIEVPAYOCCPY-----GMCASFKASGQWEADHDDESSKRPGLLAROQ 288
Db 268 LIDATLTPSHCCAFRLNPTKEQNFSEIFK-----NFSQCESTARP-----NN 313
QY 289 ENHY-----DODDELQLEMEDSKPHPSVOCSPPTPGCFKCEYLFESWGIRLAVMAIVLL 343
Db 314 ETLISALPASELSMDYDYGFCSP-KTIGCAPBPDAFNPCEIDIMGYDLRLVILINLITL 372
QY 344 SVLCNGVLTLTPRAGPAPLPVKFVVGAIAGANTLTGISCGLASVDALTFGQFSEYGA 403
Db 373 AIMGANTVLEVLTLTSHYKILVP-REFLNCNLSFADFCMGILYLLILASVADOTKGOYVNAI 431
QY 404 RMETGLGCRATGFLAVLGSEASVLLTLAAVOCSSVSCVRAVGSBSLGSVRAGVLGCL 463
Db 432 DMOTGNGCVAFETVFESELSVYTLVITLERMHTIYALDQKRLRRAJPIIMLGW 491
QY 464 ALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVAVAGAVIKL 523
Db 492 LESTILANPLVGVSSYKVKISCLPM--DVETLISQVYITLITILNVAVAIICACIYKI 549
QY 524 YCDLPRGCEAV-WDCAMVRVAMLTIPADGLVCPVAFESFASMLGFLPTPRAVAVSL 582
Db 550 YFAVONBELMATNDKTIARKKAVLFTDTCMAPISEFSAALAVPLITVNSVLLV 609
QY 583 VLELPACNPLLYLLENHFRDRLRPRAGDSGLAYAAAGELSSCDSTQA 638
Db 610 LEFVPVSCANPLIYALFTYKAFRDFPL-----LSKSGCKKHOA 648

RESULT 10
US-07-741-453A-61
Sequence 61 Application US/07741453A
Patent No. 622857
GENERAL INFORMATION:
APPLICANT: PARMENTIER, MARC
APPLICANT: LIBERT, FREDERIC
APPLICANT: DUMONT, JACOUES
APPLICANT: VASART, GILBERT
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROPHOPIN-RECEPTOR
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,453A
FILING DATE: 19911015
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N
REGISTRATION NUMBER: 16773

```

```

REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-741-453A-61

Query Match
Best Local Similarity 11.6%; Score 445; DB 4; Length 764;
Matches 167; Conservative 129; Mismatches 321; Indels 114; Gaps 20;

QY 16 QEEPAIIRTLGR-----QELGHNRIWEIGADPESQSLQALD 173
Db 35 EDRVYCKDIORIPSPSTQTLKLETHLRTPSHAFSNLPISRTIVSITLTLQDLS 94
QY 68 SAFQYLPKHTLSLNGANDIOEF-PD-LKGTSLLETLTRAGIRLPLSGMCOOLPRLR- 124
Db 95 HSFYNSKVTTHLEINRNTNLTYIDPDLKELPLKFLGIFNTGKMP-----DLTKVYS 149
QY 125 -----VLELSHN-----QIEELPSLHRCCKLEIGQHNRIWEIGADPESQSLQALD 173
Db 150 TDFILEITNPNPMTSIPVNAFGL--CNETLTKLYNNGFSVGYAFNG-TRKIDAVY 206
QY 174 LSMAA-IRSHPEAFSTLSLVK-LDLTDNLTTLPLAGLGIMHKLKGNLNAQSKDSF 231
Db 207 LNKKNKYTLVDKAFGSGVYSGPLDVSQTSVALPSKGLHKLRLANVTYTKRLPLS 266
QY 232 DSEPKRILEVPAVYOCPP-----GMCASF----- 257
Db 267 LSFHLTRADLSTPHSCAFKQKIRGILLESIMCNESMSQSLRQKSVNALNSLPQHY 326
QY 258 -----FRASQGW-----EADLHLDDESSKRPGLLAROENHYD 293
Db 327 EENLGDSTIVYKESKRFODTHNNAHYFFEEQDEITIGQELKNQOEFITLQAFSDHD 386
QY 294 QDLDELOEMEDSKPHPSVOCSPPTPGCFKCEYLFESWGIRLAVMAIVLLSVLCNGLVLL 353
Db 387 YTI-----CGDSE--DMCTPKSDEFNPCEIDIMGYKFLRYVMVSLALGAVFVL 437
QY 354 TVFAGPAPLPVKFVVGAIAGANTLTGISCGLASVDALTFGQFSEYGARMETGAGRA 413
Db 438 -ILTSYKLNVRFLMNCNLAFAFCMGWYLLILASVDLYTHSEYVNAIDMOTSPGNT 496
QY 414 TGPLAVLGSEASVLLTLAAVOCSSVSCVRAVGSBSLGSVRAGVLGCLALAGLAALP 473
Db 497 AGFTVFASELVYTLVITLERMVAITFMALDKRIKLRAAAMVWGWCCLLALLP 556
QY 474 LASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCFLVAVAGAVIKLPLPRGDFE 533
Db 557 LVGISYAKVSICLPM--DTEPLALAYIMSVLVNVAIVYVOCCKVITVVRNPOYN 614
QY 534 -AVMDCAMVRVAMLTIPADGLVCPVAFESFASMLGFLPTPRAVAVSLVLPPLACIN 592
Db 615 PDDKOTKIAKMAVILFTDFCMAPISFVALMINKNPLITVNSITLVLFLPLNSCAN 674
QY 593 PLVTLNRPHERDRLRPRAGDSGLAYAAAG-----ELEKSSQDSQALVA 641
Db 675 PELVAITTKAFORVDFITLSKFCICKRQAOATRGQVPPKRNSTDIQOVKVTIDMGOGLHN 734
QY 642 FSDVDTLLEAS 652
Db 735 MEDVYELEINS 745

```

```

RESULT 11
US-07-757-342D-5

```

```

: Sequence 5, Application US/07757342D
: Patent No. 6218509
: GENERAL INFORMATION:
: APPLICANT: IGARASHI, Masao
:             MINEGISHI, Takashi
:             NAKAMURA, Kazuto
: TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
:             CUSHMAN
: STREET: 130 Water Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/757,342D
: FILING DATE: 10-Sep-1991
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: BUCKLEY, Linda M.
: REGISTRATION NUMBER: 31003
: REFERENCE/DOCKET NUMBER: 41226
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)523-3400
: TELEFAX: (617)523-6440
: TELEX: 200291 STRE UR
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 764 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-07-757-342D-5

Query Match      11.5%; Score 444; DB 4; Length 764;
Best Local Similarity 22.8%; Pred. NO.3.8e-31;
Matches 167; Conservative 128; Mismatches 322; Indels 114; Gaps 20;

```

```

QY 354 TVPAGGAPLPVPEKVVGAAGANTLTGISCGLASVDALTFGGFSEYGARWETGLGCR 413
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 - ILTSHYKLNVPRELMCNLAFADFCMGMYLLIASVDLYTHSEYNNADIMQTPGNT 496
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 TGPLAVGSEASVLLTLAAVQCSVSCVAAYGKSPSLGSRGVLGCLALAGLAALP 473
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 497 AGFTYFASLSVTLVTLLTLEMYAITFAMRLDRKMLRHACAIMGVWCFFLLALP 556
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 LASVGEASPLCLPYAPREGOPALGFTVALVMNSCFIVVGAATKLCODLPGRDPE 533
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 557 LVGISIAKYSICLPM--DIETPLALAIYVLLNLNIAFYVCCGYKIIITYANPOYN 614
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 534 -AVWDCAMVRHVAWLIFADGLLYCPVAFLSPASMLGFPTPEAVKSVLLVPLPACIN 592
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 615 PGDKDTIAKRAVLIPTDFICMAPISFYALSAILNKPLTVNSKILLVLFYPLNSCAN 674
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 593 PLIYLFPNPHRDRLRLRRAGDSGLAYAAAG-----ELKSSDSTQALYA 641
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 PELYALFTKEFRDVFLLSKFGICKRQADYRGQRPVPKNSTDIQVKYVHEMRGLHN 734
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 642 FSDVDLLEAS 652
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 735 MEDVYELIEKS 745
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-07-757-342D-3
: Sequence 3, Application US/07757342D
: Patent No. 6218509
: GENERAL INFORMATION:
: APPLICANT: IGARASHI, Masao
:             MINEGISHI, Takashi
:             NAKAMURA, Kazuto
: TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
:             CUSHMAN
: STREET: 130 Water Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/757,342D
: FILING DATE: 10-Sep-1991
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: BUCKLEY, Linda M.
: REGISTRATION NUMBER: 31003
: REFERENCE/DOCKET NUMBER: 41226
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)523-3400
: TELEFAX: (617)523-6440
: TELEX: 200291 STRE UR
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 700 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-07-757-342D-3

Query Match      11.5%; Score 443.5; DB 4; Length 700;
Best Local Similarity 22.9%; Pred. NO.3.7e-31;
Matches 160; Conservative 117; Mismatches 300; Indels 121; Gaps 18;

```


OY 642 FSDVDILEAS 652
|| :||
Db 735 MEDVEYELIENS 745

RESULT 14
US-07-757-342D-10

; Sequence 10, Application US/07757342D
; Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 674 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-07-757-342D-10

Query Match 10.7%; Score 413; DB 4; Length 674;

Best Local Similarity 23.8%; Pred. No. 1.9e-28;

Matches 147; Conservative 109; Mismatches 287; Indels 74; Gaps 15;

OY 38 IKAIEKAMG-NPLLOTIHFYDNPLOFGRSAFOYLPRKHTLSLNGAMDIQEPFDLKG 96

Db 36 VAVTSQARGLNEVYKIEISQIDSLERIEANAFDNL-LNLS-----LHRCQK 144

OY 97 TSEILITLRAGIRLPSGMCQOLPRRLVLELSHNOIEELPS-----LHRCQK 144

Db 77 ---ELIQNTKMLRYEPGAFILPGKLYLSICNTGIRKRPDVTKYVSSSNTFLICDN 133

OY 145 LE-----EIGLOHNRIMEIGADTFSQSSLSQALDSMN-ITSIHP 186

Db 134 LHITTPGNAFOGMNESVYTLKLYGNGFEVQSHAFNG-TTLTSLLEKENVHIERKMG 192

OY 187 FSTLSLVLDLTNDQUTLPLAGLGLMLKLGMLATLQAFSKDSFPKRLILEVYAY 246

Db 193 FKGATGPKTLIDISSTYKQLPSTGLSIOPLATSSSLAKLPSTRETFVNLLEKTLTPS 252

OY 247 OCCPYGMCASFKAQWEADLHLDESSKRPGLGLAQAEHN--YDQDLDELQLEME 304

Db 253 HCCA-----FNLPTKEQNFHSISENFSKO-CESTVVRKVSNTLYSSMLAESELSGW 304

OY 305 DSK-----PHSVCCSPRPGRPKCEYLFESMGIRLAVMAITVLSLNGVLTLPVAGG 359

Db 305 DYXGFCILP-KTPRCADPEPDFAFNPCEDIMGYDFKRLITLNLITLADGN-MTVLEVLTS 362

OY 360 PAPLPYKFEVVGAIAGANTLTGISGLLASVDALFFQSEYGAEMETGLGCRATFLAV 419

Db 363 RYKLTVPFRLMCNLSFADFQMGVLLILIASVDSQTKGYNNHAIIDMQSGSGTGFFV 422

OY 420 LGSFASVLLTLAAVQSVSVSVYRAGKSPSLGSRVAGVGLCALAGIAALPLASGE 479

Db 423 FASLSVYTLTVTLERHHTLTVAIHLDQKRLRHAIIIMLGWLFSSLIAMIPLVGVSN 482

OY 480 YGASPLCLPYAPREGQRALGFYALVMMNSFCPLVAGAVYIKIYCDLPBGDEAV-MDC 538

Db 483 YMKVSIQFPM--DVETTLISQYIILITLILNVAFELICACYIKITFAVNPDLMAFNKDT 540

OY 539 AMVRHVAMLIPADGLLYCPVAFSLFASMLGLFPVPEAVKSVLYLVLPACINPLLYLL 598

Db 541 KIAKKMAILITDFTQCMAPISFPAISAPFKVPLITVNSKVLVLFYPIINSCANPELVAI 600

OY 599 FNPFRDRLRLRPAG 615

Db 601 FTKTFQRDFEFLLSKFG 617

RESULT 15
US-07-757-342D-2

; Sequence 2, Application US/07757342D
; Patent No. 6218509

GENERAL INFORMATION:

APPLICANT: IGARASHI, Masao

MINEGISHI, Takashi

NAKAMURA, Kazuo

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

STREET: 130 Water Street

CITY: Boston

STATE: Massachusetts

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,342D

FILING DATE: 10-Sep-1991

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BUCKLEY, Linda M.

REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 41226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)523-3400

TELEFAX: (617)523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 699 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-07-757-342D-2

Query Match 10.7%; Score 413; DB 4; Length 699;

Best Local Similarity 23.8%; Pred. No. 2e-28;

Matches 147; Conservative 109; Mismatches 287; Indels 74; Gaps 15;

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 disorder -

XX PS Claim 9; Fig 8; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, cancer-related wasting,
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC obsessive-compulsive disorder, anxiety, panic disorder, schizophrenia,
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a human LGR6 polypeptide.

XX Sequence 736 AA;

Query Match Best Local Similarity 100.0%; Score 3850; DB 23; Length 736;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLHLETLNLVNYNKLQEPVPAIRLGRLOELGFHNNNKAIPEKFMKNPLLOTIHFND 60
 DB 1 GLHLETLNLVNYNKLQEPVPAIRLGRLOELGFHNNNKAIPEKFMKNPLLOTIHFND 60
 QY 61 PLOVGRSAFQYLRKLTSLNGANDIOEPDLKGTSLLETLTRAGIRLLPSGMCQQL 120
 DB 61 PLOVGRSAFQYLRKLTSLNGANDIOEPDLKGTSLLETLTRAGIRLLPSGMCQQL 120
 QY 121 PRFLVLELHNQIELPSLHRCQKLEIGLOHNRIMWEGATFQSLSLQALDLSMAIR 180
 DB 121 PRFLVLELHNQIELPSLHRCQKLEIGLOHNRIMWEGATFQSLSLQALDLSMAIR 180
 QY 121 PRFLVLELHNQIELPSLHRCQKLEIGLOHNRIMWEGATFQSLSLQALDLSMAIR 180
 DB 121 PRFLVLELHNQIELPSLHRCQKLEIGLOHNRIMWEGATFQSLSLQALDLSMAIR 180
 QY 181 SHPEASTLHSLVKKLDTNOLTTPLAGIGLMLHLKGNLALSOAFSKDSFPKRLIL 240
 DB 181 SHPEASTLHSLVKKLDTNOLTTPLAGIGLMLHLKGNLALSOAFSKDSFPKRLIL 240
 QY 241 EYVYAYQCCYGCASFPGKASGQWEADLHLDDESSKRLGLARQAEHHYDDDELQ 300
 DB 241 EYVYAYQCCYGCASFPGKASGQWEADLHLDDESSKRLGLARQAEHHYDDDELQ 300
 QY 301 LEWEDSKPHPSVOCSPTPGPFKCEYLFEESWGIRLAWAIVLVLSCNGVLTLTFVAGP 360
 DB 301 LEWEDSKPHPSVOCSPTPGPFKCEYLFEESWGIRLAWAIVLVLSCNGVLTLTFVAGP 360
 QY 361 APPLPKFVAGATAGANTLTGTCGILLASVDALTFCOFSYTGARWETGLCCRAATGLAVL 420
 DB 361 APPLPKFVAGATAGANTLTGTCGILLASVDALTFCOFSYTGARWETGLCCRAATGLAVL 420
 QY 421 GSEASVLLTLAAVQCSVSCVRAGKSPISGVSAGVAGVLCIALGLAALPLASVGEY 480
 DB 421 GSEASVLLTLAAVQCSVSCVRAGKSPISGVSAGVAGVLCIALGLAALPLASVGEY 480
 QY 481 GASPLCLPAPREGOPALGFTVALVMNNSFCFLVAVGAYIKLYCDLPRGDPEAVWDCAM 540
 DB 481 GASPLCLPAPREGOPALGFTVALVMNNSFCFLVAVGAYIKLYCDLPRGDPEAVWDCAM 540
 QY 541 VRHVAMLIFFADGLLYCPVAFILSFASMLGLFPVTPPEAVKSVLLVLPPLPACLNPLLYLFFN 600

DB 541 VRHVAMLIFFADGLLYCPVAFILSFASMLGLFPVTPPEAVKSVLLVLPPLPACLNPLLYLFFN 600
 QY 601 PHFRDDLRLRPRAGDSGLAYAAAGELKSSCDSTQALVAFSDVDLILEASEGRPG 660
 DB 601 PHFRDDLRLRPRAGDSGLAYAAAGELKSSCDSTQALVAFSDVDLILEASEGRPG 660
 QY 661 ETVGFPSVTLISCOOPAPRLEGSHCVPEEGNHNHGNQPSMDGLLLRABGSPAGGGLS 720
 DB 661 ETVGFPSVTLISCOOPAPRLEGSHCVPEEGNHNHGNQPSMDGLLLRABGSPAGGGLS 720
 QY 721 GGGGFPSPGLAFASHV 736
 DB 721 GGGGFPSPGLAFASHV 736

RESULT 2

AA66141
 ID AAG6141 standard; Protein; 967 AA.

AC AAG6141;

DT 13-MAR-2002 (first entry)

DE Human LGR6 polypeptide (clone Fb150881).

KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;

KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;

KW osteopathic; antiartherosclerotic; vasotropic; hypotensive; cardiac;

KW antirhythmic; anorectic; gene therapy; human.

OS Homo sapiens.

PN WO200185768-A2.

PD 15-NOV-2001.

PF 08-MAY-2001; 2001WO-0515002.

PR 08-MAY-2000; 2000US-0566588.

PS (MILL-) MILLENNIUM PHARM INC.

PT Gu W;

DR WPI: 2002-055584/07.

DR N-PSDB: AA167926, AA167927.

PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful
 PT for diagnosing and treating weight disorder, metabolic disorder,
 PT central nervous system disorder, endocrine disorder and cardiovascular
 disorder -

XX PS Claim 9; Fig 15; 198pp; English.

CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, cancer-related wasting,
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC obsessive-compulsive disorder, anxiety, panic disorder, schizophrenia,
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection

assays (e.g., chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics), and in methods of treatment (e.g., therapeutic and prophylactic). The present sequence represents a human IgH6 polypeptide.

Sequence 967 AA;

Query Match 100.0%; Score 3850; DB 23; Length 967;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 736; Conservative 0;

1 GLHNETLDLVNKLQEPVPAIRTLGRLOELGFHNNIKAIPEKAFKGNPLDITTFYDN 60
 232 GLHNETLDLVNKLQEPVPAIRTLGRLOELGFHNNIKAIPEKAFKGNPLDITTFYDN 291
 61 PLOFVGRSAFOYLPKRLTLNLGAMDIQEPDLKGTSLLEITLRAGIRLLPSGMCOOL 120
 292 PLOFVGRSAFOYLPKRLTLNLGAMDIQEPDLKGTSLLEITLRAGIRLLPSGMCOOL 351
 121 PRLRYLESHNOIELPSLHRCQKLEELGLOHNRIMEIGADFFSOLSLQALDLSMAIR 180
 352 PRLRYLESHNOIELPSLHRCQKLEELGLOHNRIMEIGADFFSOLSLQALDLSMAIR 411
 181 SIHPEFSTLSLVKLDLTDNLTPLAGLGLMHLKGLMALSLQAFSKDSPKRLTL 240
 412 SIHPEFSTLSLVKLDLTDNLTPLAGLGLMHLKGLMALSLQAFSKDSPKRLTL 471
 241 EYVAYVQCCPYGMCASFRRKASGQWAEEDLHDESSKRPGLLAROANHHDDODELO 300
 472 EYVAYVQCCPYGMCASFRRKASGQWAEEDLHDESSKRPGLLAROANHHDDODELO 531
 301 LEMESKRPSPVQCPPTGPKFCEYLFESMGIRLVAIVLTVLVCNGLVLTVPAGGP 360
 532 LEMESKRPSPVQCPPTGPKFCEYLFESMGIRLVAIVLTVLVCNGLVLTVPAGGP 591
 361 ABLPVKFEVGAIVAGANTLTGISCGLASVDALTGQFSEYGAEMETGICRATGFLAVL 420
 592 ABLPVKFEVGAIVAGANTLTGISCGLASVDALTGQFSEYGAEMETGICRATGFLAVL 651
 421 GSEASVLTLLTAQVCSVSCVRANKSPISGVRAVLGCLALAGLAALPLASVGEY 480
 652 GSEASVLTLLTAQVCSVSCVRANKSPISGVRAVLGCLALAGLAALPLASVGEY 711
 712 GASPLCLPAPPEBGPALGFTVALVMNNSFCFLVYAGAYIKLYCDLPRGDFEAVWDCAM 771
 541 VRRVAVLITADGLLYCPVAFPLSFASMLGLFPVTPEAVKSVLVVLPPLACLNPLTLLEN 600
 772 VRRVAVLITADGLLYCPVAFPLSFASMLGLFPVTPEAVKSVLVVLPPLACLNPLTLLEN 831
 601 PPHRDRLRRLRPAGSGPLAVAAAGELEKSSCDSTQALVAFSDVLLILEASAGRPGL 660
 832 PPHRDRLRRLRPAGSGPLAVAAAGELEKSSCDSTQALVAFSDVLLILEASAGRPGL 891
 661 EYVGFPSVTLISQCGAPRLLEGSHCEVEPGNHFGNPPSMDGELLIRAGSTPAGGGLS 720
 892 EYVGFPSVTLISQCGAPRLLEGSHCEVEPGNHFGNPPSMDGELLIRAGSTPAGGGLS 951
 721 GGGGFPSPGSLAFASHV 736
 952 GGGGFPSPGSLAFASHV 967

RESULT 3
 AAE23418
 ID AAE23418 standard; Protein; 948 AA.

XX AAE23418;
 XX AC
 XX DT 27-AUG-2002 (first entry)
 XX

Human G-protein coupled receptor-10 (GCRG-10) protein.

Human; G-protein coupled receptor-10; GCRG-10; atherosclerosis; cancer; cell proliferative disorder; gastrointestinal; autoimmune; metabolic; neurological; inflammatory; cardiovascular; viral infection; anorexia; cirrhosis; multiple sclerosis; Huntington's disease; gene therapy; Alzheimer's disease; heart failure; hepatitis; cholecystitis; obesity; rheumatoid arthritis; Hashimoto's thyroiditis; angina pectoris; osteoporosis.

Homo sapiens.

Key Location/Qualifiers
 Peptide 1..25
 Protein /label= signal_peptide 26..948

Domain /note= "Mature GCRG-10 protein"

Domain /note= "Leucine rich repeat N-terminal domain"

Domain /label= Transmembrane_domain
 Domain /label= Transmembrane_domain

MO200226825-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30661.

29-SEP-2000; 2000US-236546P.

13-OCT-2000; 2000US-240589P.

20-OCT-2000; 2000US-242233P.

03-NOV-2000; 2000US-242322P.

03-NOV-2000; 2000US-245855P.

09-NOV-2000; 2000US-247587P.

15-NOV-2000; 2000US-249343P.

(INCY-) INCYTE GENOMICS INC.

Baughn MR, Graul RC, Walla NK, Gandhi AR, Hafalia AJA, Ramkumar J; Tribouley CM, Thornton M, Kalliock DA, Yao MG, Elliott VS; Burford N, Khan FA, Yue H, Lu Y, Arvizu C, Roopa R, Nguyen DB; Lee EA, Lu DM, Ison CH, Walsh RT, Policky JL; MPI: 2002-426012/45.

N-PSDB: AAD37674.

Novel G-protein coupled receptor polypeptides referred as GCRG peptides, useful for treating multiple sclerosis, cholecystitis heart failure, angina pectoris, rheumatoid arthritis, obesity, osteoporosis

Claim 71; Page 126-129; 147pp; English.

The invention relates to human G-protein coupled receptor (GCRG 1-16) polypeptides and polynucleotides. Sequences of the invention are useful for the diagnosis, treatment and prevention of cell proliferative (e.g., cancer, cirrhosis), neurological (e.g., multiple sclerosis, Huntington's disease, Alzheimer's disease), cardiovascular (e.g., angina pectoris, heart failure), gastrointestinal (e.g., anorexia, cholecystitis), autoimmune/inflammatory (rheumatoid arthritis, Hashimoto's thyroiditis) and metabolic disorders (e.g., obesity, osteoporosis), viral infections, atherosclerosis and hepatitis. GCRG proteins are useful for identifying compounds that modulate, mimic and block olfactory and taste sensations. They are also useful for identifying GCRG modulators. GCRG DNAs are useful in gene therapy. The present sequence is human GCRG-10 protein.

Query Match 99.6%; Score 3835; DB 23; Length 948;
 Best Local Similarity 99.7%; Pred. No. 0;

Matches 734; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GLHNLETLDTLNKYOEPVAIRTLGRLOELGFHNNNTKAIPEKAFMGNPLLOTTHFYDN 60
 DB 213 GLHNLETLDTLNKYOEPVAIRTLGRLOELGFHNNNTKAIPEKAFMGNPLLOTTHFYDN 272
 QY 61 PLOFVGRSAFOYLPKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSMCOOL 120
 DB 273 TIOFVGRSAFOYLPKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSMCOOL 332
 QY 121 PRLVLELSHNOIEELPSLRCKLEETIGLOHNRIMWEGADTFQSLSLQALDLSWNAIR 180
 DB 333 PRLVLELSHNOIEELPSLRCKLEETIGLOHNRIMWEGADTFQSLSLQALDLSWNAIR 392
 QY 181 SIHPEAFSTLSLVKLDLTNDQTLTTLPLAGIGLMLKLNKLNLSQAFSDSPFKRLTL 240
 DB 393 SIHPEAFSTLSLVKLDLTNDQTLTTLPLAGIGLMLKLNKLNLSQAFSDSPFKRLTL 452
 QY 241 EYPAAYOCCPYGMCASFEEKASGQWEADLHLDDESSKRPGLLAROENHYDQDLDELQ 300
 DB 453 EYPAAYOCCPYGMCASFEEKASGQWEADLHLDDESSKRPGLLAROENHYDQDLDELQ 512
 QY 301 LEMEDSKPHPSVOCSPPTPGPFKECYLTFESWGIRLAWAIVLVLNGVLLTFVAGGP 360
 DB 513 LEMEDSKPHPSVOCSPPTPGPFKECYLTFESWGIRLAWAIVLVLNGVLLTFVAGGP 572
 QY 361 ADLPVKFVVGALAGANTLTGISCGLLASVDALTFFGQSEYGARWETGLGRATGFLAVL 420
 DB 573 ADLPVKFVVGALAGANTLTGISCGLLASVDALTFFGQSEYGARWETGLGRATGFLAVL 632
 QY 421 GSEASVLLTLTAAVOCSSVSCVRAVGKSPSLGSRAGVLCIALAGLAAALPLASVGEY 480
 DB 633 GSEASVLLTLTAAVOCSSVSCVRAVGKSPSLGSRAGVLCIALAGLAAALPLASVGEY 692
 QY 481 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 540
 DB 693 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 752
 QY 541 VRHVAMLIFFADGLLYCPVAFLSFASMLGLEPVTPEAVKSVLLVVLPLPACLNPLLYLLEN 600
 DB 753 VRHVAMLIFFADGLLYCPVAFLSFASMLGLEPVTPEAVKSVLLVVLPLPACLNPLLYLLEN 812
 QY 601 PHEFRODLRLRPRAGSGGLAYAAAGELKSSCDSTQALVAFSDVLLTEASAGRPGL 660
 DB 813 PHEFRODLRLRPRAGSGGLAYAAAGELKSSCDSTQALVAFSDVLLTEASAGRPGL 872
 QY 661 ETYGPPSVTLISCOOPGAPRLGSHCVPEEGNHFGNPOPSMDGELLIRAGSTPAGGGLS 720
 DB 873 ETYGPPSVTLISCOOPGAPRLGSHCVPEEGNHFGNPOPSMDGELLIRAGSTPAGGGLS 932
 QY 721 GGGGPOPSGLAFASHV 736
 DB 933 GGGGPOPSGLAFASHV 948

RESULT 4
 AA079941
 ID AA079941 standard; Protein: 940 AA.
 AC AA079941;
 XX 15-JUL-2002 (first entry)
 DE Human orphan GPCR-like protein NOV3.
 XX Human orphan GPCR-like protein NOV3.
 KW Human; NOV3 polypeptide; cardiomyopathy; atherosclerosis; cancer;
 KW cell signal processing; metabolic pathway modulation; cancerous tissue;
 KW antibody; diabetes; transgenic animal; orphan GPCR-like protein; NOV3.
 OS Homo sapiens.
 XX
 XX
 PN WO200229038-A2.
 XX

PD 11-APR-2002.
 XX
 XX 04-OCT-2001; 2001WO-US31377.
 XX
 XX 04-OCT-2000; 2000US-237862P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Herrmann JL, Rastelli L, Shinkets RA;
 XX
 DR WPI; 2002-340104/37.
 DR N-PSDB; Abk49424.
 XX
 PT Novel isolated NOV3 polypeptide, and encoded polynucleotide, useful for
 PT treating cardiomyopathy, atherosclerosis, and cancer
 XX
 PS Claim 1; Page 36; 180pp; English.

The present invention relates to a new NOV3 polypeptide having a
 CC 900 (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or
 CC 331 (NOV6) residue amino acid sequence, as given in the specification.
 CC The novel polypeptide, and its encoding polynucleotide, are used to
 CC treat cardiomyopathy, atherosclerosis, cancer or a disease related to
 CC cell signal processing and metabolic pathway modulation, in a human.
 CC detecting the polypeptide or polynucleotide is useful for identifying
 CC cancerous tissue. The antibody can be used to treat diabetes or cancer.
 CC The host cells can be used to produce non-human transgenic animals
 CC useful in drug screening. The present amino acid sequence represents
 CC the human orphan GPCR-like protein NOV3 of the invention.
 CC
 XX
 S0 Sequence 940 AA;

Query Match 98.6%; Score 3795; DB 23; Length 940;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GLHNLETLDTLNKYOEPVAIRTLGRLOELGFHNNNTKAIPEKAFMGNPLLOTTHFYDN 60
 DB 186 GLHNLETLDTLNKYOEPVAIRTLGRLOELGFHNNNTKAIPEKAFMGNPLLOTTHFYDN 245
 QY 61 PLOFVGRSAFOYLPKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSMCOOL 120
 DB 246 PLOFVGRSAFOYLPKLTHTLSLNGAMDIOEPDLKGTTSLEILTLTRAGIRLLPSMCOOL 305
 QY 121 PRLVLELSHNOIEELPSLRCKLEETIGLOHNRIMWEGADTFQSLSLQALDLSWNAIR 180
 DB 306 PRLVLELSHNOIEELPSLRCKLEETIGLOHNRIMWEGADTFQSLSLQALDLSWNAIR 365
 QY 181 SIHPEAFSTLSLVKLDLTNDQTLTTLPLAGIGLMLKLNKLNLSQAFSDSPFKRLTL 240
 DB 366 SIHPEAFSTLSLVKLDLTNDQTLTTLPLAGIGLMLKLNKLNLSQAFSDSPFKRLTL 425
 QY 241 EYPAAYOCCPYGMCASFEEKASGQWEADLHLDDESSKRPGLLAROENHYDQDLDELQ 300
 DB 426 EYPAAYOCCPYGMCASFEEKASGQWEADLHLDDESSKRPGLLAROENHYDQDLDELQ 485
 QY 301 LEMEDSKPHPSVOCSPPTPGPFKECYLTFESWGIRLAWAIVLVLNGVLLTFVAGGP 360
 DB 486 LEMEDSKPHPSVOCSPPTPGPFKECYLTFESWGIRLAWAIVLVLNGVLLTFVAGGP 545
 QY 361 ADLPVKFVVGALAGANTLTGISCGLLASVDALTFFGQSEYGARWETGLGRATGFLAVL 420
 DB 546 ADLPVKFVVGALAGANTLTGISCGLLASVDALTFFGQSEYGARWETGLGRATGFLAVL 605
 QY 421 GSEASVLLTLTAAVOCSSVSCVRAVGKSPSLGSRAGVLCIALAGLAAALPLASVGEY 480
 DB 606 GSEASVLLTLTAAVOCSSVSCVRAVGKSPSLGSRAGVLCIALAGLAAALPLASVGEY 665
 QY 481 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 540
 DB 666 GASPLCLPYAPPEGOPALGFTVALVMNSFCFLVAGATIKLYCDLPRGDFEAVWDCAM 725
 QY 541 VRHVAMLIFFADGLLYCPVAFLSFASMLGLEPVTPEAVKSVLLVVLPLPACLNPLLYLLEN 600

```

|||||
Db 726 VRHVAWMLFEADGLLYCPVAFLEFASMLGLFVPTPEAVKSVLLVLPAPACLNPLLYLTFN 785
QY 601 PHEHDDLRRLPRAGDSGLPVAAGLELEKSSCDSTQALVAESVDVLLIEASAGRPGI 660
Db 786 PHEHDDLRRLPRAGDSGLPVAAGLELEKSSCDSTQALVAESVDVLLIEASAGRPGI 845
QY 661 ETYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQSMDELLLRAGSTPAGGGLS 720
Db 846 ETYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQSMDELLLRAGSTPAGGGLS 905
QY 721 GGGGFP 727
Db 906 GGAFAFP 912

RESULT 5
AAV53575
ID AAV53575 standard; Protein; 847 AA.
AC AAV53575;
XX 15-FEB-2000 (first entry)
DE Human gonadotropin receptor partial sequence #5.
XX Human gonadotropin receptor; screening; ligand; biomedical research;
KW biochemical research; drug; hormone; reproductive tissue; infertility;
KW contraception.
XX Homo sapiens.
XX EP950711-A2.
XX 20-OCT-1999.
XX PD 02-FEB-1999; 99EP-0200303.
XX PF 06-FEB-1998; 98EP-0200357.
XX PR 27-JUL-1998; 98EP-0202519.
XX PR 24-SEP-1998; 98EP-0203213.
XX PA (ALKU) AKZO NOBEL NV.
XX Van Der Spek PJ, Heikoop JC;
DR WPI: 1999-563673/48.
DR N-PSDB: AAZ40461.
PT New 7 transmembrane gonadotropin receptors, useful for screening for
PT hormone analogs and drugs -
PS Claim 1: Page 24-28; 38pp; English.
CC Sequences AAV53571-Y53578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.
CC XX
SQ Sequence 847 AA:
Query Match 97.8%; Score 3767; DB 20; Length 847;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 720; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 LNHLETLIDINYNKLOEPVAVIRTLGRLOELGFHNNNIKAIPEKAFMGNPLOTIHFYDNP 61
I : : : : :

```

```

Db 94 LPSTQSLDLYNKNLOEFPVAVIRTLGRLOELGFHNNNIKAIPEKAFMGNPLOTIHFYDNP 153
QY 62 IOEFGSAFOYLPKLTHTLSLNGAMDIOEPDLKGTSTSEILITLRAGIRLLPSMGCOQLP 121
Db 154 IOEFGSAFOYLPKLTHTLSLNGAMDIOEPDLKGTSTSEILITLRAGIRLLPSMGCOQLP 213
QY 122 RLRLVLESHNOIEELPSLRQKLEFELQHNRIWEIGADTFSSQSLQALDLSMNAIIS 181
Db 214 RLRLVLESHNOIEELPSLRQKLEFELQHNRIWEIGADTFSSQSLQALDLSMNAIIS 273
QY 182 IHPEAFSTLHSLVKLDLTDNOLTTPLPLAGLGLMHLKLGKLNALSOAFSEKDSFPKLRILE 241
Db 274 IHPEAFSTLHSLVKLDLTDNOLTTPLPLAGLGLMHLKLGKLNALSOAFSEKDSFPKLRILE 333
QY 242 VPYAYQCCPYGMCASFRRASQWEADLHLDDESSKRPGLGLARQENHYQDDLELO 301
Db 334 VPYAYQCCPYGMCASFRRASQWEADLHLDDESSKRPGLGLARQENHYQDDLELO 393
QY 302 EMEBSKPHPSVQCSPTPPGPFKCEYLFESWGIIRLAWAIVLLSVLCNGLVLLTFVAGGPA 361
Db 394 EMEBSKPHPSVQCSPTPPGPFKCEYLFESWGIIRLAWAIVLLSVLCNGLVLLTFVAGGPA 453
QY 362 PLPPYKFEVGAIAAGANTLTGISCGLASVDALTFQGFSEYGARWETGLCRATGLAVLG 421
Db 454 PLPPYKFEVGAIAAGANTLTGISCGLASVDALTFQGFSEYGARWETGLCRATGLAVLG 513
QY 422 SEASVLLTLTAAVQCSVSVCAVATGKSPSLGSVRAGVIGCLALAGLAALPLASVGEYG 481
Db 514 SEASVLLTLTAAVQCSVSVCAVATGKSPSLGSVRAGVIGCLALAGLAALPLASVGEYG 573
QY 482 ASPICLPYAPPEGOPAAIGFTALVMMNSFCFLVAVAGATIKYICDILPRGDFEAVMCAW 541
Db 574 ASPICLPYAPPEGOPAAIGFTALVMMNSFCFLVAVAGATIKYICDILPRGDFEAVMCAW 633
QY 542 RHVAMLIFADGLLYCPVAFLEFASMLGLFVPTPEAVKSVLLVLPAPACLNPLLYLTFN 601
Db 634 RHVAMLIFADGLLYCPVAFLEFASMLGLFVPTPEAVKSVLLVLPAPACLNPLLYLTFN 693
QY 602 HFRDRLRLRPRAGDSGLPVAAGLELEKSSCDSTQALVAESVDVLLIEASAGRPGLE 661
Db 694 HFRDRLRLRPRAGDSGLPVAAGLELEKSSCDSTQALVAESVDVLLIEASAGRPGLE 753
QY 662 TYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQSMDELLLRAGSTPAGGGLS 721
Db 754 TYGFPSTVTLISCOOPGAPRLGSHCVPEEGNHFGNPPQSMDELLLRAGSTPAGGGLS 813
QY 722 GGGGFP 727
Db 814 GGAFAFP 819

RESULT 6
AAG66138
ID AAG66138 standard; Protein; 967 AA.
AC AAG66138;
XX 13-MAR-2002 (first entry)
DE Mouse IGR6 polypeptide (clone ftmzb048h10).
XX Large G-protein coupled receptor 6; IGR6; GPCR; immunomodulator;
KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antiangine;
KW osteopathic; antiarteriosclerotic; vasodilator; hypotensive; cardiant;
KW antiarrhythmic; anorectic; gene therapy; mouse.
XX XX
OS Mus musculus.
XX Key 1.23 Location/Qualifiers
FH Peptide /note="signal peptide"
FT Protein 24..967

```

/note= "mature protein"

MO200185768-A2.

15-NOV-2001.

08-MAY-2001; 2001MO-US15002.

08-MAY-2000; 2000US-0566588.

(MILL-) MILLENNIUM PHARM INC.

Gu W:

WPI: 2002-055584/07.

N-PSDB: AA167920, AA167921.

Novel isolated large G-protein coupled receptor 6 polypeptide, useful for diagnosing and treating weight disorder, metabolic disorder, central nervous system disorder, endocrine disorder and cardiovascular disorder

Example 1; Fig 1; 198pp; English.

The invention provides isolated large G-protein coupled receptor 6 (LGR6) polypeptides and polynucleotides encoding them. The LGR6 polypeptides and polynucleotides are useful as targets for developing modulating agents that regulate a variety of cellular process, e.g. neural and endocrine processes, as well as thermogenesis. They are useful for developing novel diagnostic and therapeutic agents for LGR6 associated disorders such as weight disorders (anorexia, obesity), eating disorders, acquired immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting, metabolic disorders, central nervous system disorders (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia, obsessive-compulsive disorder, anxiety, panic disorder, migraine), endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism, acromegaly), and cardiovascular disorders (atherosclerosis, restenosis, hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides, polynucleotides and antibodies are useful in screening assays, detection assays (e.g., chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics), and in methods of treatment (e.g. therapeutic and prophylactic). The present sequence represents a mouse LGR6 polypeptide.

Sequence 967 AA:

Query Match 88.9%; Score 3423; DB 23; Length 967;

Best Local Similarity 89.3%; Pred. No. 5.1e-298;

Matches 657; Conservative 31; Mismatches 48; Indels 0; Gaps 0;

```

OY 1  G L H N E T T L D L N Y N K L O E F P A I R T L G L G F H N N N I K A I P E K A F N G P L I O T I H Y E Y D N 60
DB 232 G L H N E T T L D L N Y N K L O E F P A I R T L G L G F H N N N I K A I P E K A F N G P L I O T I H Y E Y D N 291
OY 61  P I O F P G R S A F O Y L P K U L H T S L N G A M D I O E F P D I K G T T S E I L T T R A G I R L L P S G M C O O L 120
DB 292 P I O F P G R S A F O Y L S K L H T S L N G A T D I O E F P D I K G T T S E I L T T R A G I R L L P P G V C O L 351
OY 121 P R I R V L E T S H N O E E L P S L H R C K L E I G I O H R I W E I G A D T S Q L S S I O A L D S M A I R 180
DB 352 P R I R I E L S H N O E E L P S L H R C K L E I G I R H R I K E I G A D T S Q L S I O A L D S M A I R 411
OY 181 S I H P E A F S T L H S L V K L D L D N Q L T T L P L A G I G L G I M H L K L G N L A L S Q A F S K D S P K R I L 240
DB 412 A I H P E A F S T L R S L V K L D L D N Q L T T L P L A G I G L G I M H L K L G N L A L S Q A F S K D S P K R I L 471
OY 241 E V P Y A Y Q C C P Y G M C A S F F R A S G M E A D L H L D E S S K R P L G I L A R O A E N H Y D D I D E L Q 300
DB 472 E V P Y A Y Q C C A Y G I C A S F F T S G M O A E D F H P E E E A P K R P L G I L A G O A E N H Y D I D E L Q 531

```

```

OY 301 L E M E D S K P H P S V O C S P P T P P E F K P C E Y L P E S N G I T L A Y A I V L S L C N G I V L I V P A G C P 360
DB 532 M G T E D S K P N P S V O C S P P G E F K P C E H L P E S W G I R L A W A I V L S L C N G I V L I V F A S G P 591
OY 361 A P L P Y K V E V G A I G A N T L G I S C G L L A S V D A L T F E G G E S E G A M E T G L G C R A G F L A V L 420
DB 592 S P L S P Y K V L V G M A G A N A L G I S C G L L A S V D A L T Y G G F A E I G A M E S G L G C A T G F L A V L 651
OY 421 G S E A S V L L T L A A V O C S V S V C V R A Y K S P S L G S V R A G I G C I A L A G I A A L P L A S V G E Y 480
DB 652 G S E A S V L L T L A A V O C S I S T V C V R A Y K A P S P G S V R A G I G C I A L A G I A A L P L A S V G E Y 711
OY 481 G A S P L C P A P P E G P A A L G E T V A L V M N S E C F L Y A G A T I K L C D P P R G F E A V M D C A M 540
DB 712 G A S P L C P T A P P G R P A L S F A V A L V M N S L C F L V A G A T I K L C D L P R G D F E A V M D C A M 771
OY 541 V R H V A M L I F A D G L L Y C P V A F L S F A S M I G L P V P P E A V K S V L V L P P A C I N P L I L Y L F N 600
DB 772 V R H V A M L I F A D G L L Y C P V A F L S F A S M I G L P V P P E A V K S V L V L P P A C I N P L I L Y L F N 831
OY 601 P H R D D L R L R P R A G S G P L A Y A A G E L E K S C D S T Q A L V A F S D V D L I L E A S E A G R P P G L 660
DB 832 P H R D D L R L R L P S P R A P G P L A Y A A G E L E K S C D S T Q A L V A F S D V D L I L E A S E A G R P P G L 891
OY 661 E T Y G F S V T L I S C Q O G A P R L E G S H C V E P E G N H F G N P O P M D G E L L I R A G S P P A G G L S 720
DB 892 E T Y G F S V T L I S H Q G A I R L E G N H F I E S D G T F G N P O P M K G E L L K A G A T L A G C G S S 951
OY 721 G G G F O P S G I A F A S H V 736
DB 952 V G G A L M P S G S L F A S H L 967

RESULT 7
AAG66139
ID AAG66139 standard; Protein; 633 AA.
AC AAG66139;
XX
DT 13-MAR-2002 (first entry)
XX
DE Human LGR6 polypeptide (clone fah).
XX
KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antiangine;
KW osteopathic; antiatherosclerotic; vasotropic; hypotensive; cardiac;
KW antiarrhythmic; anorectic; gene therapy; human.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FT Misc-difference 57
FT /label= unknown
FT /note= "encoded by CTN"
XX
XX WO200185768-A2.
XX
XX 15-NOV-2001.
XX
XX 08-MAY-2001; 2001MO-US15002.
XX
XX 08-MAY-2000; 2000US-0566588.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Gu W:
XX
XX WPI: 2002-055584/07.
XX
XX N-PSDB: AA167922, AA167923.
XX
XX Novel isolated large G-protein coupled receptor 6 polypeptide, useful
XX for diagnosing and treating weight disorder, metabolic disorder,
XX

```

PT central nervous system disorder, endocrine disorder and cardiovascular
PT disorder -
PS
PS Example 1; Fig 5; 198pp; English.

XX The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC polynucleotides are useful as targets for developing modulating agents and
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes, as well as thermogenesis. They are useful for developing novel
CC diagnostic and therapeutic agents for LGR6 associated disorders such as
CC weight disorders (anorexia, obesity), eating disorders, acquired
CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, Gilles de la
CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC polynucleotides and antibodies are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials and pharmacogenomics), and in methods of
CC treatment (e.g. therapeutic and prophylactic). The present sequence
CC represents a human LGR6 polypeptide.

XX Sequence 633 AA;

Query Match 78.3%; Score 3015; DB 23; Length 633;
Best Local Similarity 99.3%; Pred. No. 1.2e-261;
Matches 576; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 157 EIGADTFSSLSIQALDLSWNAIRSIHPEAFSTLSLVKLDLTDNQLTTPPLAGIGLMH 216
DB :: |||||
54 KLEXDFFSOLSIQALDLSWNAIRSIHPEAFSTLSLVKLDLTDNQLTTPPLAGIGLMH 113
OY 217 LKLGKGLALSQAASKSPFKRLTEVPYAYAOCCPYGMCASFRRKSGWEMEDLHDEES 276
DB |||||
114 LKLGKGLALSQAASKSPFKRLTEVPYAYAOCCPYGMCASFRRKSGWEMEDLHDEES 173
OY 277 SKRPGLILARQENHNDODLEDELEMEDESKPHPSVOCSTTPGPFKCEYFEESMGIRLA 336
DB |||||
174 SKRPGLILARQENHNDODLEDELEMEDESKPHPSVOCSTTPGPFKCEYFEESMGIRLA 233
OY 337 WVAIVLLSVLCNGVLVLTFFAGGAPAPLPVKFVVGAIAGANTLIGISGILLASVDALTFG 396
DB |||||
234 WVAIVLLSVLCNGVLVLTFFAGGAPAPLPVKFVVGAIAGANTLIGISGILLASVDALTFG 293
OY 397 QPSEVGARWETGIGCRATGFLAVLGEASVLLTLTAVOCSSVSVSCVRANGKSPSLGSVR 456
DB |||||
294 QPSEVGARWETGIGCRATGFLAVLGEASVLLTLTAVOCSSVSVSCVRANGKSPSLGSVR 353
OY 457 AGVIGCLIALAGLAAALPLASVGEYASPLCLPYAPPEGOPAAAGFTVALVMNSFCFLVY 516
DB |||||
354 AGVIGCLIALAGLAAALPLASVGEYASPLCLPYAPPEGOPAAAGFTVALVMNSFCFLVY 413
OY 517 AGAYIKLYCDLPGRDFEAVWDCAMVRHVAWLIFADGILLYCPVAFSLFASMLGLPVPYPEA 576
DB |||||
414 AGAYIKLYCDLPGRDFEAVWDCAMVRHVAWLIFADGILLYCPVAFSLFASMLGLPVPYPEA 473
OY 577 VASVLLVLPPLPACLNPLLYLTFNPHFRDDLRLRPAGSGPLATVAAAGLEKSSCDST 636
DB |||||
474 VASVLLVLPPLPACLNPLLYLTFNPHFRDDLRLRPAGSGPLATVAAAGLEKSSCDST 533
OY 637 QALVAFSDVDLLEASGAPPLGLEYGPPSVTLISQOOPGAPLLESHCHCEPBNHNGN 696
DB |||||
534 QALVAFSDVDLLEASGAPPLGLEYGPPSVTLISQOOPGAPLLESHCHCEPBNHNGN 593
OY 697 POPSMDELLLRAGSTPAGGGLSGGGFQPSGLAFASHV 736
DB |||||
594 POPSMDELLLRAGSTPAGGGLSGGGFQPSGLAFASHV 633

RESULT 8
ABG09314
ID ABG09314 standard; Protein; 691 AA.

XX
AC ABG09314;
DT 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #9305.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
OS
PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HSE-) HXSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS73501.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 39673; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 691 AA;

Query Match 75.0%; Score 2886.5; DB 22; Length 691;
Best Local Similarity 89.1%; Pred. No. 4.8e-250;
Matches 580; Conservative 10; Mismatches 48; Indels 13; Gaps 9;

OY 85 MDIOEPFDIKGTTSLEILTLTRAGIRLPLPSGKQQLPRLVLELSHNQIEPLSLARCK 144

DB 1 MDIOEPFDIKGTTSLEILTLTRAGIRLPLPSGKQQLPRLVLELSHNQIEPLSLARCK 60

OY 145 LEEIGIQRHRIEIGADTFSSLSIQALDLSWNAIRSIHPEAFSTLSLVKLDLTDNQLT 204

```

Db 61 LEEIGLQHRRIWEIGADFTSOLSLQALDLSMNAIRSIHBEAFSTLSHVLDTTSQLT 120
QY 205 TLPLAGLGLMHLKTKGNALSOAFSKDSPKRLILEVPAYOCSPYGMCSFRRASQW 264
Db 121 TLPLAGLGLMHLKTKGNALSOAFSKDSPKRLILEVPAYOCSPYGMCSFRRASQW 180
QY 265 EAEDLHLDDESSKRPLGLARQAEHNYDQDLDELQLEMEDSKRHPVQCSPTGPKPC 324
Db 181 EAEDLHLDDESSKRPLGLARQAEHNYDQDLDELQLEMEDSKRHPVQCSPTGPKPC 240
QY 325 EYLFESWGIIRLAVMAIVLLSVLCNGLVLTFRAGPAPLPYKFFVGAAGANTLTGISC 384
Db 241 EYLFESWGIIRLAVMAIVLLSVLCNGLVLTFRAGPAPLPYKFFVGAAGANTLTGISC 300
QY 385 GLLASVDALTFGQFSEYRG-ARHMETGLGCRATGFLAVYGEASVLLTLTAAYQCSVASCV 443
Db 301 GLLASVDALTFGQFSEYRGARHMETGLGCRATGFLAVYGEASVLLTLTAAYQCSVASCV 360
QY 444 RAYGKPSLGSVRAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGOPALGFTV 503
Db 361 RAYGKPSLGSVRAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGOPALGFTV 420
QY 504 ALVMMNSCFLVVAAG-YYIKLYCDLPFGDFEAVWDCAMVRHV-AN-LIFADGLICPV 558
Db 421 ALVMMKLLLFSEWVGPLPYHQMNYCDLPFGDFEAVWDCAMVRHVANLHKTGLICPV 480
QY 559 AFLSFASMLGLFPVTPPEAVKSVLLVPLPACLNPLLYLTFNPHFRDRLRLR-PRAGDS 617
Db 481 AFLSFASMLGLFPVTPPEAVKSVLLVPLPACLNPLLYLTFNPHFRDRLRLRPRRRTQ 540
QY 618 GPLVYAAAGEL--EKSQDS-QTALVAFSDVDL-ILEASEA-GRPGLETYGFPSVTLI 671
Db 541 GPPXPFLCCGREDEGASCDYSYQALVAVPLSMISFECAXSMAVGLLETFYGFPSVTLI 600
QY 672 SCQOGCARLESCHVEPEGNHFGNPPQPSMDGELLRLRAESTPAGGSLSG 722
Db 601 SCQOGCARLESTHCEVEPEGNHFGNPPQPSMDGELLRLRAESTPAGGSLSG 651

RESULT 9
AA53574
ID AA53574 standard; Protein; 497 AA.
XX
AC AA53574;
XX
DT 15-FEB-2000 (first entry)
XX
DE Human gonadotropin receptor partial sequence #4.
XX
KW Human; gonadotropin receptor; screening; ligand; biomedical research;
KW biochemical research; drug; hormone; reproductive tissue; infertility;
KW contraception.
XX
OS Homo sapiens.
XX
PN EP950711-A2.
XX
PD 20-OCT-1999.
XX
PF 02-FEB-1999; 99EP-0200303.
XX
PR 06-FEB-1998; 98EP-0200357.
PR 27-JUL-1998; 98EP-0202519.
PR 24-SEP-1998; 98EP-0203213.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Van Der Spek PJ, Heikoop JC;
XX
DR WPI; 1999-563673/48.
XX
DR N-PSDB; AA240460.
XX

```

```

PT New 7 transmembrane gonadotropin receptors, useful for screening for
PT hormone analogs and drugs -
XX
PS Claim 1; Page 18-21; 38pp; English.
XX
CC Sequences AA53571-Y3578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.
XX
SQ Sequence 497 AA:

Query Match 63.6%; Score 2450; DB 20; Length 497;
Best Local Similarity 99.6%; Pred. No. 4,7e-211;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 259 KASGWEAEDLHLDDESSKRPLGLARQAEHNYDQDLDELQLEMEDSKRHPVQCSPTP 318
Db 1 KASGWEAEDLHLDDESSKRPLGLARQAEHNYDQDLDELQLEMEDSKRHPVQCSPTP 60
QY 319 GPPKPEYLFESWGIIRLAVMAIVLLSVLCNGLVLTFRAGPAPLPYKFFVGAAGANT 378
Db 61 GPPKPEYLFESWGIIRLAVMAIVLLSVLCNGLVLTFRAGPAPLPYKFFVGAAGANT 120
QY 379 LTGISGGLASVDALTFGQFSEYGARHMETGLGCRATGFLAVYGEASVLLTLTAAYQCSV 438
Db 121 LTGISGGLASVDALTFGQFSEYGARHMETGLGCRATGFLAVYGEASVLLTLTAAYQCSV 180
QY 439 SVSCVARYGKPSLGSVRAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGOPAA 498
Db 181 SVSCVARYGKPSLGSVRAGVGLGCLALAGLAALPLASVGEYASPLCLPYAPPEGOPAA 240
QY 499 LGFTVALVMMNSCFLVVAAGAYIKLYCDLPFGDFEAVWDCAMVRHVANLIFADGLICPV 558
Db 241 LGFTVALVMMNSCFLVVAAGAYIKLYCDLPFGDFEAVWDCAMVRHVANLIFADGLICPV 300
QY 559 AFLSFASMLGLFPVTPPEAVKSVLLVPLPACLNPLLYLTFNPHFRDRLRLRPRAGDSG 618
Db 301 AFLSFASMLGLFPVTPPEAVKSVLLVPLPACLNPLLYLTFNPHFRDRLRLRPRAGDSG 360
QY 619 PLAYAAAGELKSSCDSTOALVAFSDVDLILEASEAGRPPELETYGFPSVTLISQOPGA 678
Db 361 PLAYAAAGELKSSCDSTOALVAFSDVDLILEASEAGRPPELETYGFPSVTLISQOPGA 420
QY 679 PRLEGSHCEVEPEGNHFGNPPQPSMDGELLRLRAESTPAGGSLSGGQFOP 727
Db 421 PRLEGSHCEVEPEGNHFGNPPQPSMDGELLRLRAESTPAGGSLSGGQFOP 469

RESULT 10
AAG67556
ID AAG67556 standard; Protein; 395 AA.
XX
AC AAG67556;
XX
DT 26-NOV-2001 (first entry)
XX
DE A human guanosine triphosphate binding protein coupled receptor.
XX
KW Human; guanosine triphosphate binding protein coupled receptor;
KW G protein coupled receptor; brain disease; cancer.
XX
OS Homo sapiens.
XX
PI WO200109323-A1.
XX
DR WO200109323-A1.
XX
DR 08-FEB-2001.
XX

```

[illegible]

KW extracellular leucine rich repeat region; mapping; identification.
 XX Homo sapiens.
 OS
 PN WO9948921-A1.
 XX
 PD 30-SEP-1999.
 XX
 PE 25-MAR-1999; 99WO-US06573.
 XX
 PR 26-MAR-1998; 98US-0079501.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX (ORCA) ORGANON NV.
 PI Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
 XX WPI; 1999-591074/50.
 XX DR N-PSDB; AAZ25344.
 XX
 PT New G-protein coupled receptors, useful for identifying their own
 PT ligands -
 PS
 PS Claim 2; Fig 2; 54pp; English.
 XX
 CC The present sequence represents the human G-protein coupled receptor
 CC having extracellular leucine rich repeat regions, designated LGR5.
 CC The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
 CC the receptor. The polypeptides and/or polynucleotides are also useful
 CC for homologous or related genes, producing compositions that modulate
 CC the expression or function of the receptors, gene therapy, mapping
 CC functional regions of the receptors, studying associated physiological
 CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
 CC for producing antibodies, and for identifying biologically active
 CC agents. The polypeptides contain a G-protein coupled seven
 CC transmembrane region and a leucine rich repeat extracellular domain.
 CC These regions capture and facilitate optimal orientation of its ligand.
 CC The proteins are also expressed in diverse tissues.
 CC
 SQ Sequence 693 AA:
 Query Match 47.8%; Score 1842; DB 20; Length 693;
 Best Local Similarity 52.2%; Pred. No. 3.1e-156;
 Matches 360; Conservative 113; Mismatches 186; Indels 30; Gaps 8;
 1 GLHNETLIDLNKLOEFPVARTLIGRLQELGPHNNNIKAIPKAFKGNPLLOTIHEYDN 60
 18 GLHSLTETLDNKNNDLDEFPARTLSMLKELGPHNNIRSIPEKAFGNPSLTIHFYDN 77
 61 PLOFVGRSAFOYLPKLTHTLSNGAMDIQERPDKGTSLLELTITRAGIRLLPSGMCOOL 120
 78 PLOFVGRSAFOHLPKLTHTLNGASQITTEPPDLGTANLESLITGAQISSLPQVCNOL 137
 121 PRLRVLELSHNOEELPSLRCKLEIEIGLOHNRWEIGATFQSLSIALDLDSWNAIR 180
 138 PNLQVLDLSLNLLEDLPSEFVCKQKIDLRHNEIYEIKDTFOQLLSRLNLAMNNTIA 197
 181 SHPEAFSTLHSLVKLDLTNQLTTPLAGLGLMLHLKLGKGNLALSQAFSDSPFKELT 240
 198 IHPNMFSTLPSLIKIDLSNLLSPFITGLHGLTLKLTGNHALQSWISSENEPELKV 257
 241 EYVYAVOCCYGCASFVKASGOW-----EADPLHLDDESSRKPLGLLRQAQNHVDQ 294
 258 EMPYAYOCCAFGCENYKISNQMNKGNSSMDLH-----KKDAGMFOAQDE----R 306
 295 LDDELQLEM-EDSKPHPSVOCSPPTGPFKPCLEYLFESNGIRLAWAIVLVLCGLVLL 353
 307 LDDELDFEDLKLHLSVOCSPSPGPFKPCENHLIDGLIRIGVWTIAVLATCNALVTS 366
 354 TVRAGGPAPLPVKFVVGALAGANTLTIGISGLASVDALTFGQSFEGARRETLGCR 413
 367 TVRR-SPLYTSPITKLITGLVIAAVNMMLTGVSSAVLAGVDAFTFGSPARRHGAMWENGCHV 425

OY 414 TGFVLAVGSEASVLLLTIAVQCVSVSCVARYCKSPSLGVSRAVGLGIALAGIAALP 473
 DB 426 IGFSLIFASESSVFLTLAALERGEFSVYSAKFEKAPFSSILKLTLLCALLALMAAVP 485
 OY 474 LASVGEYASPLCLPYAPPEGOPALGFTVALVMNNSCFVLVAVAGATIKLYCDLPRGDFE 533
 DB 486 LLGSKRTGASPLCLPL--PGEPSFTGIMVALILNLSLCFLLMTITATYKLYCNLDKGDLE 543
 OY 534 AVMDCAWRYHVAWMLIFADGLIYCPVAFISFASMLGLFPVTPKAVKSVLLVLPPLACINP 593
 DB 544 NIMDCSMWKHIALLLFTNCILNCPVAFISFSSLLNLTFSPEVIFILLVVPPLACINP 603
 OY 594 LLYLLENPHFDLRLRPRA-----GDSGPLATYAAAGLEKSSGDSRQALVAESDVDL 648
 DB 604 LLYILFNPHFEKEDLVSLRKQTYWTRSKHPSLMSINSDDVERKSDSTOALVTFTSSIT 663
 OY 649 LEASEAGRP---PGLFTYGFSPVTLISC 673
 DB 664 YDLPPSSVSPRAYPYTESCHLSVAFPVC 692

RESULT 14

AA90687
 ID AAY90687 standard; Protein; 907 AA.
 AC
 XX AAY90687;
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human mutant G protein-coupled receptor HG38 (V765K).
 DE
 XX
 KW G protein-coupled receptor; GPCR; constitutively active;
 KW intracellular loop 3; transmembrane domain 6; drug screening;
 KW agonist; antagonist; mutant; mutein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO20002129-A1.
 XX
 XX 20-APR-2000.
 PD
 XX
 PF 12-OCT-1999; 99WO-US23938.
 XX
 PR 13-OCT-1998; 98US-0170496.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Chalmers DT, Liaw CW;
 XX
 DR WPI; 2000-329165/28.
 DR N-PSDB; AAA30779.
 XX
 PT Non-endogenous constitutively activated human G protein-coupled
 PT receptors, useful for identifying agonists for use as pharmaceutical
 PT agents -
 XX

Example 2; Page 332-335; 341pp; English.

The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643- AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743 and AAA30775-A30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X, is substituted for an endogenous residue in IC3 at a position 16 amino acids N-terminal of an endogenous proline in TM6 to form a sequence X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or Ala, and is preferably Lys. When the endogenous residue at this position is Lys, this residue is replaced by His, Arg or preferably Ala. The 15 amino acid stretch between the substituted amino acid and the Pro may be endogenous, non-endogenous, or a mixture of endogenous and non-endogenous residues. The constitutively active GPCRs are useful for

CC identifying antagonists, agonists and partial agonists for use as
 CC pharmaceutical agents. The mutant proteins are also useful in research
 CC settings for elucidating the roles of the receptors in normal and
 CC diseased conditions. Antagonists for a particular GPCR are useful for
 CC treating diseases and disorders associated with that receptor. Because
 CC the novel mutant GPCRs are constitutively active, they can be used
 CC directly for screening of compounds without the need for endogenous
 CC ligands. Sequences AAY90643 - AAY90677 and AAY90683-Y90687 the mutant
 CC human GPCRs of the invention.

XX Sequence 907 AA;

Query Match 47.7%; Score 1836; DB 21; Length 907;
 Best Local Similarity 52.1%; Pred. No. 1.e-155;
 Matches 359; Conservative 113; Mismatches 187; Indels 30; Gaps 8;

QY 1 GLHNLETLIDLVNKKLOEPPVAIRTLGRLQELGPHNNNIKAPEKAFMGNPILLQTHIFPDN 60
 DB 232 GLHSLETLIDLVNKKLOEPPVAIRTLGRLQELGPHNNNIKAPEKAFMGNPILLQTHIFPDN 291
 QY 61 PIQVGRSAFOYLPKLTLSLNGAMDIOEFPDLKGTSLLETLITRAGIRLLPSGMCQOL 120
 DB 292 PIQVGRSAFOYLPKLTLSLNGASQTEFPDLGTANLESILTGAQISSLPOTVCNQL 351
 QY 121 PRLRVLESHNOIEELPSLHRCCKLEETIGLOHNRWEIGADTESLSLQALDLSMNAIR 180
 DB 352 PRLRVLESHNOIEELPSLHRCCKLEETIGLOHNRWEIGADTESLSLQALDLSMNAIR 411
 QY 181 SHPEAFSTHSLVRLDNLQTLPLAGLGIMHLKLGKLNALSAFSDSPKRLIL 240
 DB 412 IHPNFASTLPSIKRLDNLSSLSLSPITGLHGTILKLGKLNALQSLISENPELKI 471
 QY 241 EYVAYOCCPYGMCASFKAAGOW-----EADLHLDESSKRPGLLARAQENHYDQ 294
 DB 472 EYVAYOCCPYGMCASFKAAGOW-----EADLHLDESSKRPGLLARAQENHYDQ 520
 QY 295 DLDELQLEK-EDSKRHPVOCSTPBPFPKCEYLFESMGIRLAWAIVLLSVLCNGLVLL 353
 DB 521 DLDELQLEK-EDSKRHPVOCSTPBPFPKCEYLFESMGIRLAWAIVLLSVLCNGLVLL 580
 QY 354 TVEAGGAPALPVRKPVVGAIGANLTLTGISCGLASVADALTFGOFSEYGARWETGLCCRA 413
 DB 581 TVEAGGAPALPVRKPVVGAIGANLTLTGISCGLASVADALTFGOFSEYGARWETGLCCRA 639
 QY 414 TGFVLAVSEASVLLTLTAAVQCSVSVSCVRAVYKSPSLGSRAGVLCGLALAGLAALP 473
 DB 640 TGFVLAVSEASVLLTLTAAVQCSVSVSCVRAVYKSPSLGSRAGVLCGLALAGLAALP 699
 QY 474 LASVEYGASPLCLPYAPPEGPPALGFTVALVMMNSFCFLVAGAVYIKLXCDLPRGDFE 533
 DB 700 LASVEYGASPLCLPYAPPEGPPALGFTVALVMMNSFCFLVAGAVYIKLXCDLPRGDFE 757
 QY 534 AVMDCAVVRHVAWMLTFADOLLYCPAFLSPASMLGLEPVPPEAVKSVLLVLEPLACINP 593
 DB 758 AVMDCAVVRHVAWMLTFADOLLYCPAFLSPASMLGLEPVPPEAVKSVLLVLEPLACINP 817
 QY 594 LLYLLEHFRDRLRLRPRP-----GDSGRLAYAAAGLEBSSCDSTOALYAFSPVDLI 648
 DB 818 LLYLLEHFRDRLRLRPRP-----GDSGRLAYAAAGLEBSSCDSTOALYAFSPVDLI 877
 QY 649 LEASEAGRP-----PGLETYGFPSVTLISC 673
 DB 878 YDLPPSSVPSFAYPTESCHLSSVAFVPC 906

RESULT 15

AAE22835 ID AAE22835 standard; Protein: 951 AA.

XX AAE22835;

DT 21-AUG-2002 (first entry)

XX

DE Human AXOR33 protein.

KW Human; 7 transmembrane receptor; AXOR33 protein; infection; pain; cancer;
 KW diabetes; neurological disease; Parkinson's disease; depression; stroke;
 KW delirium; dementia; cardiovascular disease; hypertension; gene therapy;
 KW vaccine; acute heart failure; protein therapy; analgesic; antimicrobial;
 KW cytostatic; neuroprotective; cerebroprotective; cardiovascular;
 KW neotropic.

XX Homo sapiens.

PN WO200153337-A1.

XX 26-JUL-2001.

PF 12-JAN-2001; 2001WO-US01051.

PR 20-JAN-2000; 2000US-0488354.

PA (SMIK) SMITHKLINE BEECHAM CORP.

XX Elshourbagy NA, Gattu M, Shabon U, Taylor AH;

DR WPI; 2001-465365/50.

XX N-PSDB; AAD36904.

PT New human 7 transmembrane receptor AXOR33 polypeptide and
 PT polynucleotide, useful for identifying agonists or antagonists useful
 PT in therapy, as well as diagnosing or treating e.g. infections, pain,
 PT cancer or neurological diseases
 XX Claim 1; Page 26-27; 36pp; English.

CC The present invention relates to human 7 transmembrane receptor AXOR33
 CC proteins and polynucleotides encoding such proteins. Sequences of the
 CC invention are useful for diagnosing or treating certain diseases such
 CC as infections (e.g. bacterial, viral or fungal), pain, cancer, diabetes,
 CC neurological diseases (e.g. Parkinson's disease, depression, delirium,
 CC dementia or stroke) or cardiovascular disease (e.g. acute heart failure
 CC or hypertension). AXOR33 sequences are used as vaccines. They are also
 CC used in gene therapy and protein therapy. The present sequence is human
 CC AXOR33 protein.

XX Sequence 951 AA;

Query Match 36.9%; Score 1420; DB 22; Length 951;
 Best Local Similarity 43.0%; Pred. No. 4e-118;
 Matches 302; Conservative 105; Mismatches 236; Indels 60; Gaps 9;

QY 1 GLHNLETLIDLVNKKLOEPPVAIRTLGRLQELGPHNNNIKAPEKAFMGNPILLQTHIFPDN 60
 DB 223 GLDNLLETDLVNNLGEPPQAIKALPSLKELGPHNSISVITDGAFDGPFLLRTIHLIDN 282
 QY 61 PIQVGRSAFOYLPKLTLSLNGAMDIOEFPDLKGTSLLETLITRAGIRLLPSGMCQOL 120
 DB 283 PIQVGRSAFOYLPKLTLSLNGAMDIOEFPDLKGTSLLETLITRAGIRLLPSGMCQOL 342
 QY 121 PRLRVLESHNOIEELPSLHRCCKLEETIGLOHNRWEIGADTESLSLQALDLSMNAIR 180
 DB 343 KMLFTLIDLVNKKLOEPPVAIRTLGRLQELGPHNNNIKAPEKAFMGNPILLQTHIFPDN 402
 QY 181 SHPEAFSTHSLVRLDNLQTLPLAGLGIMHLKLGKLNALSAFSDSPKRLIL 240
 DB 403 EIHSHRAFATLGPITNLDVSEFNEISFPEGLNGLNQKLVDGNFKELAKDPFNALSL 462
 QY 241 EYVAYOCCPYGMCASFKAAGOW-----EADLHLDESSKRPGLLARAQENHYDQ 296
 DB 463 EYVAYOCCPYGMCASFKAAGOW-----EADLHLDESSKRPGLLARAQENHYDQ 518
 QY 297 DELQLEDESKRHPVOCSTPBPFPKCEYLFESMGIRLAWAIVLLSVLCNGLVLL 356
 DB 519 -----IINCTPSTGAFCFCEYLLGSMWIRLIVWFTFLVALFRLNLVILTFE 564

